

Inverse Scale Space Decomposition

Marie Foged Schmidt¹, Martin Benning², Carola-Bibiane Schönlieb²

¹ Technical University of Denmark, Department of Applied Mathematics and Computer Science, Artusvej 5, Bldg. 303B, 2800 Kongens Lyngby, Denmark

² University of Cambridge, Department of Applied Mathematics and Theoretical Physics, Wilberforce Road, Cambridge CB3 0WA, United Kingdom

E-mail: mfs@dtu.dk, mb941@cam.ac.uk, cbs31@cam.ac.uk

Abstract. We investigate the inverse scale space flow as a decomposition method for decomposing data into generalised singular vectors. We show that the inverse scale space flow, based on convex and absolutely one-homogeneous regularisation functionals, can decompose data represented by the application of a forward operator to a linear combination of generalised singular vectors into its individual singular vectors. We verify that for this decomposition to hold true, two additional conditions on the singular vectors are sufficient: orthogonality in the data space and inclusion of partial sums of the subgradients of the singular vectors in the subdifferential of the regularisation functional at zero.

We also address the converse question of when the inverse scale space flow returns a generalised singular vector given that the initial data is arbitrary (and therefore not necessarily in the range of the forward operator). We prove that the inverse scale space flow is guaranteed to return a singular vector if the data satisfies a novel dual singular vector condition.

We conclude the paper with numerical results that validate the theoretical results and that demonstrate the importance of the additional conditions required to guarantee the decomposition result.

Key words: Generalised Singular Vectors, Inverse Scale Space Flow, Singular Value Decomposition, Source Conditions, Non-linear Spectral Transform, Total Variation Regularisation, Compressed Sensing.

Submitted to: *Inverse Problems*

1. Introduction

Regularisation methods are essential tools for the stable approximation of solutions of ill-posed inverse problems. Hence, their analysis is a vital part of inverse problems research. For linear regularisation methods a quite complete theory based on singular value decomposition has been available for a while (cf. [1, 2]). Only recently though a more general definition of singular vectors for non-linear regularisation methods has been formalised in [3], despite many previous and recent works on generalised Eigenfunctions in the context of non-linear partial differential equations and functional inequalities (cf. [4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14]), control theory (cf. [15]), image processing (cf. [16, 17, 18, 19, 20, 21, 22]), and machine learning (cf. [23, 24]). The generalised singular vector definition is based on the minimisation of convex, one-homogeneous but not necessarily differentiable functionals. Classical non-linear regularisation methods for which generalised singular vectors apply include the famous total variation (TV) regularisation model [25] and ℓ^1 -norm regularisation, which is a key tool in compressed sensing (cf. [26, 27, 28, 29]). Newer advancements include the total generalised variation (TGV, cf. [30, 31, 32, 33, 34]), regularisations based on the ℓ^1 -norm in combination with the shearlet transform [35], vectorial total variation regularisation (cf. [36, 37, 38]), structure-tensor extensions of the total variation (cf. [39, 40, 41, 42, 43]), vector-field regularisation [44], de-biased total variation-type regularisations [45] and regularisations based on the ℓ^1 -norm in combination with tight wavelet frames [46, 47, 48], to name just a few. It has been shown that generalised singular vectors play a vital role in the understanding of those non-linear regularisation methods similar to the role of classical singular vectors for linear regularisations. Key results are that the non-linear regularisation methods are capable of recovering individual singular vectors from data given in terms of a forward model applied to a generalised singular vector and additional measurement errors, just with a systematic bias. Considering corresponding inverse scale space methods (cf. [49]) as regularisation methods instead, these methods are capable of recovering generalised singular vectors without systematic bias (cf. [3]). The latter also has significant impact on discretisations of the inverse scale space method, such as Bregman iteration (cf. [50, 51]), linearised Bregman iterations (cf. [52, 53, 54, 55]) and modifications (cf. [56, 57]).

In [58] Gilboa initiated the idea of a non-linear spectral decomposition of singular vectors by defining a spectrum based on the forward-scale space formulation of the TV model. He was able to show that this non-linear spectrum can be converted into the original signal via a linear inverse transform. In [59] he made the connection that a single generalised singular vector (called Eigenfunction therein) of TV can be isolated via this spectral TV transform. In [60] the idea of the non-linear spectral transform has further been extended to arbitrary one-homogeneous functionals, and to variational regularisation as well as the inverse scale space method instead of just the forward TV flow.

The definition of the non-linear spectral transform immediately gave rise to the question of circumstances under which it would be possible to decompose a signal composed of singular vectors. In [61] this question has briefly been addressed, hinting that some sort of orthogonality condition is necessary for perfect separation (see [61, Proposition 1]). These initial ideas have been made more precise in [62]; given two singular vectors that are (fully orthogonal [62, Definition 5] and) linear in the subdifferential

[62, Definition 7], the non-linear spectral transform can separate these singular vectors perfectly (see [62, Proposition 4]). In [63] it has further been shown that the non-linear spectral transform can indeed decompose data into a finite set of singular vectors, given that the corresponding regularisation functional is a one-norm concatenated with a linear matrix such that the matrix applied to its transpose is diagonally dominant (see [63, Theorem 9]).

Despite containing impressive results, we however also observe limitations of the previously mentioned works. First of all, all research is restricted to the non-linear spectral transform, which (in its current form) is not applicable to general inverse problems. Secondly, we either have theoretical conditions that allow us to separate only two singular vectors for general one-homogeneous regularisation functionals, or we have conditions for a specific family of regularisers that allow the decomposition of data into multiple singular vectors. With this work we want to fill in the knowledge gaps by introducing two conditions that guarantee the decomposition of sums of finitely many singular vectors via the inverse scale space method. The first condition will be an orthogonality condition on the finite set of singular vectors, while the second condition will ensure that partial sums of their subgradients are contained in the subdifferential of the corresponding regularisation functional at zero. Our main contribution will be a theorem which states that perfect decomposition of a finite linear combination of generalised singular vectors is guaranteed if these two conditions are satisfied. Subsequently, we are also going to investigate the case of arbitrary data that is not given in terms of a finite linear combination of generalised singular vectors.

The paper is organised as follows. First, we are going to introduce notation and mathematical preliminaries that are necessary throughout this paper. Then we formulate and discuss the major contribution of this work which is a result that states that finite sums of generalised singular vectors can be decomposed via the inverse scale space method under suitable conditions. Subsequently we also address the question of how to characterise the first inverse scale space step in case the given data is arbitrary. In this case we investigate under which conditions the first step is a generalised singular vector. Finally, we will present numerical results to support the theoretical results and conclude with an outlook of open questions and problems.

2. Mathematical Preliminaries

In this section we set the notation, define assumptions and give mathematical preliminaries. Throughout the paper we consider inverse problems of the form

$$Ku = f \tag{2.1}$$

where $K : \mathcal{U} \rightarrow \mathcal{H}$ is a bounded linear operator from a Banach space \mathcal{U} to a Hilbert space \mathcal{H} . We will denote the space of bounded linear operators between two normed spaces \mathcal{U} and \mathcal{V} by $\mathcal{L}(\mathcal{U}, \mathcal{V})$. We denote by \mathcal{U}^* the dual Banach space of \mathcal{U} with norm

$$\|p\|_{\mathcal{U}^*} = \sup_{\|u\|_{\mathcal{U}}=1} |p(u)| = \sup_{u \in \mathcal{U} \setminus \{0\}} \frac{|p(u)|}{\|u\|_{\mathcal{U}}} = \sup_{\|u\|_{\mathcal{U}} \leq 1} |p(u)|,$$

where functional $p(u) = \langle p, u \rangle_{\mathcal{U}^* \times \mathcal{U}}$ is the dual pairing between p and u that we will abbreviate by $\langle p, u \rangle$ throughout the course of this work. If \mathcal{U} is a Hilbert space then

the dual pairing can be identified with the inner product on \mathcal{U} .

A popular variational framework for approximating solutions of (2.1) is the Tikhonov-type variational regularisation framework of the form

$$\hat{u} \in \arg \min_{u \in \text{dom}(J)} \left\{ \frac{1}{2} \|Ku - f\|_{\mathcal{H}}^2 + \alpha J(u) \right\}, \quad (2.2)$$

where $J : \text{dom}(J) \subseteq \mathcal{U} \rightarrow \mathbb{R} \cup \{+\infty\}$ is a regularisation functional that incorporates prior information about \hat{u} and $\alpha > 0$ is a regularisation parameter that controls the impact of J on \hat{u} .

Throughout this paper we assume that the regularisation functional J is proper, lower semi-continuous (l.s.c) and convex. We will further assume that J is absolutely one-homogeneous, i.e.

$$J(cu) = |c|J(u), \quad \forall c \in \mathbb{R}. \quad (2.3)$$

Such a functional is non-negative by definition and fulfils the triangle inequality

$$J(u + v) \leq J(u) + J(v). \quad (2.4)$$

Hence J is in fact a semi-norm. Furthermore, if $v \in \text{dom}(J)$ and $v_0 \in \ker(J)$ then

$$J(v + v_0) = J(v), \quad (2.5)$$

see [63, Lemma 3.2]. The subdifferential $\partial J(u)$ of J at some element $u \in \text{dom}(J)$ will be used several times throughout this work and is defined by

$$\partial J(u) = \{p \in \mathcal{U}^* : J(v) - J(u) - \langle p, v - u \rangle \geq 0, \forall v \in \text{dom}(J)\}.$$

Since J is absolutely one-homogeneous the subdifferential can be characterised by

$$\partial J(u) := \{p \in \mathcal{U}^* : \langle p, u \rangle = J(u), \langle p, v \rangle \leq J(v), \forall v \in \text{dom}(J)\}, \quad (2.6)$$

see for instance [64, Lemma 3.12]. As a special case we have

$$\partial J(0) := \{p \in \mathcal{U}^* : \langle p, v \rangle \leq J(v), \forall v \in \text{dom}(J)\}$$

since $J(0) = 0$. We observe that the subdifferential for absolutely one-homogenous functionals is invariant with respect to positive scaling, i.e.

$$\partial J(cu) = \partial J(u), \quad \forall c > 0. \quad (2.7)$$

We also want to emphasise that for a subgradient $p \in \partial J(u)$ we know by the characterisation (2.6) that $\langle p, v \rangle \leq J(v)$ for all $v \in \text{dom}(J)$. Now take $v \in \ker(J)$. Then $\langle p, v \rangle \leq 0$. Using $-v$ we then get $\langle p, v \rangle \geq 0$. This shows that for absolutely one-homogeneous functionals:

$$p \in \partial J(u) \quad \Rightarrow \quad \langle p, v \rangle = 0, \quad \forall v \in \ker(J). \quad (2.8)$$

In [50] a contrast-enhancing alternative to (2.2) named Bregman iteration has been introduced. This is an iterative regularisation method for which

$$u^{k+1} \in \arg \min_{u \in \text{dom}(J)} \left\{ \frac{1}{2} \|Ku - f\|_{\mathcal{H}}^2 + \alpha D_J^{p^k}(u, u^k) \right\},$$

where $u^0 \in \ker(J)$ and the subgradient $p^k \in \partial J(u^k)$ satisfies $p^0 \equiv 0$ and

$$p^{k+1} = p^k + \frac{1}{\alpha} K^*(f - Ku^{k+1}), \quad (2.9)$$

for all $k \in \mathbb{N}$. Here $K^* : \mathcal{H} \rightarrow \mathcal{U}^*$ denotes the Banach adjoint operator of K defined by $\langle Ku, v \rangle_{\mathcal{H}} = \langle u, K^*v \rangle_{\mathcal{U} \times \mathcal{U}^*}$ for $u \in \mathcal{U}$, $v \in \mathcal{H}$. The term $D_J^{p^k}(u^{k+1}, u^k)$ represents the (generalised) Bregman distance between u^{k+1} and u^k w.r.t. the functional J and the subgradient $p^k \in \partial J(u^k)$. The Bregman distance (cf. [65, 66]) is defined by

$$D_J^p(u, v) = J(u) - J(v) - \langle p, u - v \rangle, \quad p \in \partial J(v).$$

Note that the generalised Bregman distance is always non-negative due to the convexity of J . Since the Bregman iteration has a semi-convergence behaviour a stopping criterium for the iteration is needed.

In the limit $\alpha \rightarrow \infty$ we can interpret $\Delta t = 1/\alpha$ as a time step that tends to zero and the update (2.9) of p^{k+1} for the Bregman iteration can be interpreted as a forward Euler discretisation of the inverse scale space (ISS) flow (cf. [49])

$$\partial_t p(t) = K^*(f - Ku(t)), \quad p(t) \in \partial J(u(t)), \quad (2.10)$$

with $p(0) = 0$ and $u(0) = u^0 \in \ker(J)$. If we assume $u(0) = 0 \in \ker(J)$, from the definition of the ISS flow we then obtain $p(t) = tK^*f$ for $0 \leq t < t_1$ for some $t_1 > 0$. Hence we would require $p(t) = tK^*f \in \partial J(0)$ for $0 \leq t < t_1$, which by (2.8) implies $\langle K^*f, v \rangle = 0$ for all $v \in \ker(J)$. If this is not the case we can project K^*f onto the space of elements $p \in \mathcal{U}^*$ for which $\langle p, v \rangle = 0$ for all $v \in \ker(J)$. In the end we can always add $u^0 \in \ker(J)$ to $u(t)$. For the remainder of this work we therefore assume w.l.o.g. that $u(0) = u^0 = 0$ for the initial value of the ISS flow.

We want to point out that the solution $u(t)$ for the ISS flow has to satisfy an orthogonality condition:

Proposition 2.1 (Necessary Orthogonality Condition). *Let $u(t)$ be a solution of the inverse scale space flow (2.10). Then $u(t)$ satisfies the orthogonality condition*

$$\langle K^*(f - Ku(t)), u(t) \rangle = 0. \quad (2.11)$$

Proof. Using the chain rule for differentiation and $p(t) \in \partial J(u(t))$ we get

$$\frac{d}{dt} J(u(t)) = \langle p(t), \partial_t u(t) \rangle.$$

On the other hand, since $p(t) \in \partial J(u(t))$ and by the absolute one-homogeneity of J we get

$$\frac{d}{dt} J(u(t)) = \frac{d}{dt} \langle p(t), u(t) \rangle = \langle p(t), \partial_t u(t) \rangle + \langle \partial_t p(t), u(t) \rangle.$$

Hence we see that a solution $u(t)$ of (2.10) needs to satisfy

$$\langle \partial_t p(t), u(t) \rangle = 0.$$

From the definition of the inverse scale space flow we know $\partial_t p(t) = K^*(f - Ku(t))$ and hence

$$\langle K^*(f - Ku(t)), u(t) \rangle = 0.$$

□

If we ignore the one-homogeneity of J for a moment it becomes evident that the ISS flow is a generalisation of Showalter's method [67], i.e.

$$\partial_t u(t) = K^*(f - Ku(t)), \quad (2.12)$$

which is (2.10) for $J(u) = \frac{1}{2} \|u\|_{L^2(\Omega)}^2$, where $\Omega \subset \mathbb{R}^m$ is some domain. It is well known that for compact K between Hilbert spaces, solutions of (2.12) - as those of other linear regularisation methods - can conveniently be expressed in terms of K 's system of classical singular vectors. That is, for two orthonormal bases $\{v_j\}_{j \in \mathbb{N}}$ of $\overline{\mathcal{R}(K)}$ and $\{u_j\}_{j \in \mathbb{N}}$ of $\overline{\mathcal{R}(K^*)}$, where $\mathcal{R}(K)$ is the range of K and $\mathcal{R}(K^*)$ is the range of K^* , and a null sequence $\{\sigma_j\}_{j \in \mathbb{N}}$ satisfying

$$Ku_j = \sigma_j v_j \quad \text{and} \quad K^*v_j = \sigma_j u_j$$

for all $j \in \mathbb{N}$, we can solve (2.12) via

$$u(t) = \sum_{j=1}^{\infty} \left(1 - e^{-\sigma_j^2 t}\right) \frac{1}{\sigma_j} \langle f, v_j \rangle u_j. \quad (2.13)$$

If we now assume that our data is given in terms of some linear combination of finitely many singular vectors, i.e. we have $f = Ku^\dagger$ for

$$u^\dagger = \sum_{j=1}^n \gamma_j u_j$$

where $\gamma_j \in \mathbb{R}$ and $n \in \mathbb{N}$, then (2.13) simplifies to

$$u(t) = \sum_{j=1}^n \left(1 - e^{-\sigma_j^2 t}\right) \gamma_j u_j.$$

Hence, Showalter's method allows us to recover a weighted linear combination of the singular vectors of u^\dagger from data $f = Ku^\dagger$.

One of our main goals throughout the course of this work is to study solutions of the ISS flow (2.10). With the example of Showalter's method we have seen that for the specific choice of $J(u) = \frac{1}{2} \|u\|_{L^2(\Omega)}^2$ solutions of (2.12) are just simple transformations of the classical singular vectors if the data f is represented by a composition of those. However, as we are interested in (2.10) in combination with absolutely one-homogeneous regularisation functionals, classical singular vectors are of no use for our

study. A remedy to overcome this issue seems to be the use of so-called *generalised singular vectors* as defined in [3]. Generalised singular vectors are based on the idea of minimising the generalised Rayleigh quotient

$$u_g \in \arg \min_{\substack{u \in \ker(J)^\perp \\ \|Ku\|_{\mathcal{H}}=1}} \frac{J(u)}{\|Ku\|_{\mathcal{H}}} . \quad (2.14)$$

Here u_g denotes the so-called *non-trivial ground state* of the functional J , following [4], for which we have $J(u_g) \leq J(u)$ for all $u \in \text{dom}(J)$ with $\|Ku\|_{\mathcal{H}} = 1$. The set $\ker(J)^\perp$ is the orthogonal complement of $\ker(J)$ for which we follow the definition of [3] and define

$$\ker(J)^\perp = \{u \in \text{dom}(J) : \langle Ku, Kv \rangle = 0, \forall v \in \ker(J)\}.$$

Note that we have to ensure $\ker(K) \cap \ker(J) = \{0\}$ for this definition to make sense. Otherwise every element $u \in \ker(K) \cap \ker(J)$ would both satisfy $u \in \ker(J)$ and $u \in \ker(J)^\perp$.

We want to highlight that we require the normalisation constraint in (2.14); otherwise every function cu_g with $c \in \mathbb{R} \setminus \{0\}$ is a solution of (2.14) given that u_g is already a solution of (2.14). We could then make c arbitrarily close to zero to decrease $J(cu_g) = |c|J(u_g)$ due to the absolute one-homogeneity of J .

Now let us reformulate (2.14) in terms of the Lagrange saddle-point problem

$$(u_g, \mu_g, \eta_g) = \arg \min_u \max_\mu \max_\eta \left\{ \frac{J(u)}{\|Ku\|_{\mathcal{H}}} + \mu (\|Ku\|_{\mathcal{H}} - 1) + \eta \langle Ku, Kv \rangle \right\}, \quad (2.15)$$

with the corresponding Lagrange multipliers μ_g and η_g , for all $v \in \ker(J)$. We want to investigate the optimality condition of (2.15) with respect to u_g more closely, i.e.

$$0 \in \frac{\partial J(u_g) \|Ku_g\|_{\mathcal{H}} - \frac{J(u_g)}{\|Ku_g\|_{\mathcal{H}}} K^* Ku_g}{\|Ku_g\|_{\mathcal{H}}^2} + \frac{\mu_g}{\|Ku_g\|_{\mathcal{H}}} K^* Ku_g + \eta_g K^* Kv. \quad (2.16)$$

Computing the dual pairing of (2.16) with u_g therefore yields the relation

$$0 = \underbrace{\frac{J(u_g) \|Ku_g\|_{\mathcal{H}} - J(u_g) \|Ku_g\|_{\mathcal{H}}}{\|Ku_g\|_{\mathcal{H}}^2}}_{=0} + \underbrace{\mu_g \|Ku_g\|_{\mathcal{H}}}_{\neq 0} + \underbrace{\eta_g \langle Kv, Ku_g \rangle}_{=0},$$

which implies $\mu_g = 0$. If we further compute the dual product of (2.16) with $v \in \ker(J)$ we observe

$$0 = \underbrace{\frac{\langle p_g, v \rangle \|Ku_g\|_{\mathcal{H}} - \frac{J(u_g)}{\|Ku_g\|_{\mathcal{H}}} \langle Ku_g, Kv \rangle}{\|Ku_g\|_{\mathcal{H}}^2}}_{=0} + \frac{\mu_g}{\|Ku_g\|_{\mathcal{H}}} \underbrace{\langle Ku_g, Kv \rangle}_{=0} + \eta_g \|Kv\|_{\mathcal{H}}^2,$$

where $\langle p_g, v \rangle = 0$ for $p_g \in \partial J(u_g)$ due to (2.8). Hence, unless the kernel of J is trivial, we immediately observe $\eta_g = 0$.

Investigating (2.16) for $\mu_g = 0$, $\eta_g = 0$ and $\|Ku_g\|_{\mathcal{H}} \neq 0$ leaves us with the condition

$$\lambda_g K^* Ku_g \in \partial J(u_g), \quad (2.17)$$

for $\lambda_g := J(u_g)/\|Ku_g\|_{\mathcal{H}}^2$ and $p_g \in \partial J(u_g)$. Using the normalisation $\|Ku_g\|_{\mathcal{H}} = 1$ and due to the absolute one-homogeneity (2.3) we conclude

$$\lambda_g = \langle p_g, u_g \rangle = J(u_g).$$

Equation (2.17) now characterises the ground state that solves (2.14). However, there may exist other functions satisfying

$$\lambda K^* Ku_\lambda \in \partial J(u_\lambda), \quad (2.18)$$

with

$$\lambda = \frac{J(u_\lambda)}{\|Ku_\lambda\|_{\mathcal{H}}^2} \geq \lambda_g = \frac{J(u_g)}{\|Ku_g\|_{\mathcal{H}}^2}. \quad (2.19)$$

A function u_λ that satisfies (2.18) and (2.19) is called a *generalised singular vector* with *generalised singular value* λ . We normally assume $\|Ku_\lambda\|_{\mathcal{H}} = \|Ku_g\|_{\mathcal{H}} = 1$. Throughout the course of this work we will use the term *singular vectors* for generalised singular vectors as well. Note that the term *vector* does not refer to an element in \mathbb{R}^m but to an element in a possibly infinite dimensional space. Additionally, note that if u_λ is a singular vector then also $-u_\lambda$ is a singular vector for absolutely one-homogeneous J .

Example 2.2. We want to briefly demonstrate that the concept of generalised singular vectors (2.18) is indeed a generalisation of classical singular vectors in inverse problems. If we consider the absolutely one-homogeneous regularisation functional $J(u) = \|u\|_{L^2(\Omega)}$ for $\mathcal{U} = L^2(\Omega)$ on some domain $\Omega \subset \mathbb{R}^m$ we obtain $\partial J(u) = \{u/\|u\|_{L^2(\Omega)}\}$ for u with $\|u\|_{L^2(\Omega)} \neq 0$. Hence, (2.18) reads as

$$\lambda K^* Ku_\lambda = \frac{u_\lambda}{\|u_\lambda\|_{L^2(\Omega)}}$$

for $\lambda = \|u_\lambda\|_{L^2(\Omega)}/\|Ku_\lambda\|_{\mathcal{H}}^2$. If we multiply this equation by $1/\lambda$ and define $\sigma = \|Ku_\lambda\|_{\mathcal{H}}/\|u_\lambda\|_{L^2(\Omega)}$, we therefore obtain the classical singular vector condition

$$K^* Ku_\lambda = \sigma^2 u_\lambda.$$

Example 2.3. Let $K = I : H_0^1(\Omega) \rightarrow L_0^2(\Omega)$ be the embedding operator from $H_0^1(\Omega)$ into $L_0^2(\Omega)$, where subscript zero means that the functions are zero on the boundary of the bounded domain $\Omega \subset \mathbb{R}^m$, and let $J(u) = \|\nabla u\|_{L^2(\Omega; \mathbb{R}^m)}$. Then (2.18) reads as

$$-\frac{\|\nabla u_\lambda\|_{L^2(\Omega; \mathbb{R}^m)}^2}{\|u_\lambda\|_{L^2(\Omega)}^2} u_\lambda = \Delta u_\lambda,$$

which is simply the Eigenvalue problem of the Laplace operator.

Example 2.4. Let $K = I : \ell^1(\mathbb{R}^m) \rightarrow \ell^2(\mathbb{R}^m)$ be an embedding operator, and let $J(u) = \|Wu\|_{\ell^1(\mathbb{R}^m)}$ where $W : \ell^1(\mathbb{R}^m) \rightarrow \ell^1(\mathbb{R}^m)$ and $W^*W = I$. The subdifferential of J at $u \in \ell^1(\mathbb{R}^m)$ is characterised by

$$\partial J(u) = \left(W^* \circ \partial \|\cdot\|_{\ell^1(\mathbb{R}^m)} \right) (Wu),$$

where

$$\partial \|v\|_{\ell^1(\mathbb{R}^m)} = \text{sign}(v), \quad \text{sign}(v)_i \in \begin{cases} \{1\}, & v_i > 0 \\ \{-1\}, & v_i < 0 \\ [-1, 1], & v_i = 0 \end{cases}.$$

Then every vector $u_\lambda \in \mathbb{R}^m$ making Wu_λ consist of peaks of the same magnitude is a singular vector:

$$\lambda K^* K u_\lambda = \frac{\|Wu_\lambda\|_{\ell^1(\mathbb{R}^m)}}{\|u_\lambda\|_{\ell^2(\mathbb{R}^m)}^2} u_\lambda = \frac{\|Wu_\lambda\|_{\ell^1(\mathbb{R}^m)}}{\|u_\lambda\|_{\ell^2(\mathbb{R}^m)}^2} W^* W u_\lambda = W^* \left(\frac{\|Wu_\lambda\|_{\ell^1(\mathbb{R}^m)}}{\|u_\lambda\|_{\ell^2(\mathbb{R}^m)}^2} Wu_\lambda \right).$$

We need to show $\frac{\|Wu_\lambda\|_{\ell^1(\mathbb{R}^m)}}{\|u_\lambda\|_{\ell^2(\mathbb{R}^m)}^2} Wu_\lambda \in \text{sign}(Wu_\lambda)$. Now

$$\|u_\lambda\|_{\ell^2(\mathbb{R}^m)}^2 = \langle u_\lambda, u_\lambda \rangle = \langle u_\lambda, W^* W u_\lambda \rangle = \langle Wu_\lambda, Wu_\lambda \rangle = \|Wu_\lambda\|_{\ell^2(\mathbb{R}^m)}^2.$$

If c is the magnitude of the peaks of Wu_λ and n is the number of peaks then $\|Wu_\lambda\|_{\ell^1(\mathbb{R}^m)} = n|c|$ and $\|Wu_\lambda\|_{\ell^2(\mathbb{R}^m)}^2 = nc^2$. Hence

$$\frac{\|Wu_\lambda\|_{\ell^1(\mathbb{R}^m)}}{\|u_\lambda\|_{\ell^2(\mathbb{R}^m)}^2} Wu_\lambda = \frac{1}{|c|} Wu_\lambda \in \text{sign}(Wu_\lambda).$$

If we further want u_λ to be normalised we need $c = \pm \frac{1}{\sqrt{n}}$. The singular value λ then simplifies to $\lambda = \sqrt{n}$.

Example 2.5. Let $J(u) = \text{TV}_*(u)$ be the slightly modified total variation regularisation as defined in [3, Section 4.1]:

$$\text{TV}_*(u) := \sup_{\substack{\varphi \in C^\infty(\Omega; \mathbb{R}^m), \\ \|\varphi\|_{L^\infty(\Omega, \mathbb{R}^m)} \leq 1}} \int_{\Omega} u \operatorname{div} \varphi \, dx,$$

where $\Omega \subseteq \mathbb{R}^m$ is a bounded domain. In contrast to the normal TV-functional we do not choose test functions φ with compact support, which yields an additional boundary term. For functions $u \in W^{1,1}([0,1]) \cap C([0,1])$ we have

$$\text{TV}_*(u) = \int_0^1 |u'(x)| \, dx + |u(1)| + |u(0)|.$$

Let $K = I : BV([0,1]) \rightarrow L^2([0,1])$ be an embedding operator. It is then shown in [3, Theorem 2] that the Haar wavelet basis is an orthonormal set of singular vectors for K and J . Examples are for instance the one-dimensional Haar wavelets

$$u_{\lambda_1}(x) = \begin{cases} 1, & 0 \leq x < 1/2 \\ -1, & 1/2 \leq x < 1 \\ 0, & \text{otherwise} \end{cases}, \quad \text{and} \quad u_{\lambda_2}(x) = \begin{cases} \sqrt{2}, & 0 \leq x < 1/4 \\ -\sqrt{2}, & 1/4 \leq x < 1/2 \\ 0, & \text{otherwise} \end{cases},$$

with $\lambda_1 = \text{TV}_*(u_{\lambda_1}) = 4$ and $\lambda_2 = \text{TV}_*(u_{\lambda_2}) = 4\sqrt{2}$.

To see why generalised singular vectors are interesting in the context of inverse scale space regularisation, we recall the following theorem from [3]:

Theorem 2.6 ([3, Theorem 9]). *Let J be proper, convex, l.s.c and absolutely one-homogeneous and let u_λ be a generalised singular vector with corresponding singular value λ . Then, if the data f is given by $f = \gamma Ku_\lambda$ for a positive constant γ , a solution of the inverse scale space flow (2.10) is given by*

$$u(t) = \begin{cases} 0, & t < t_1, \\ \gamma u_\lambda, & t_1 \leq t, \end{cases}$$

for $t_1 = \lambda/\gamma$.

Hence, the ISS flow (2.10) is able to recover u_λ from data $f = Ku_\lambda$ with no systematic bias. In other words, generalised singular vectors are invariant under inverse scale space regularisation. This makes generalised singular vectors useful tools for the analysis of the ISS flow.

A second result that we want to recall is the following result in case of data $f = \gamma Ku_\lambda + g$ that can additively be decomposed into a singular vector and an arbitrary element $g \in \mathcal{H}$.

Theorem 2.7 ([3, Theorem 10]). *Let J be proper, convex, l.s.c and absolutely one-homogeneous and let u_λ be a generalised singular vector with corresponding singular value λ . Then, if the data f is given by $f = \gamma Ku_\lambda + g$ for a positive constant γ , such that there exist constants μ and η satisfying*

$$\mu K^* Ku_\lambda + \eta K^* g \in \partial J(u_\lambda) \quad (2.20)$$

with $\gamma > \mu/\eta$, a solution of the inverse scale space flow (2.10) is given by

$$u(t) = \begin{cases} 0, & t < t_1, \\ \left(\gamma + \frac{\lambda - \mu}{\eta}\right) u_\lambda, & t_1 \leq t < t_2, \end{cases}$$

for $t_1 = (\lambda\eta)/(\lambda + \gamma\eta - \mu) < t_2 = \eta$.

The original rationale behind Theorem 2.7 was to allow for deterministic noise in the data, whilst showing that the singular vector can still be recovered. However, the additional term g can obviously also be interpreted as a correction term for the singular vector u_λ to match data f .

In fact, (2.20) is nothing but a specific type of source condition. For a given function u^\dagger , the *source condition* is defined as

$$\mathcal{R}(K^*) \cap \partial J(u^\dagger) \neq \emptyset.$$

It ensures the existence of a source element $v \in \mathcal{H} \setminus \{0\}$ such that

$$K^* v \in \partial J(u^\dagger) \quad (\text{SC})$$

is satisfied. We immediately see that (2.20) is equivalent to (SC) for the specific choice $v = \mu Ku_\lambda + \eta g$ and $u^\dagger = u_\lambda$.

The (generalised) *strong source condition* as introduced in [68], in contrast to (SC), is defined as

$$\mathcal{R}(K^*K) \cap \partial J(u^\dagger) \neq \emptyset,$$

which guarantees the existence of a source element $w \in \mathcal{U} \setminus \{0\}$ with

$$K^*Kw \in \partial J(u^\dagger). \quad (\text{SSC})$$

We want to highlight the following connections between the singular vector condition (2.18) and the two source conditions (SC) and (SSC). Starting with (SSC) it is obvious that any singular vector u_λ satisfying (2.18) does also satisfy (SSC) with $u^\dagger = u_\lambda$ and $w = \lambda u_\lambda$. The converse is not true in general, which we want to demonstrate with the following example. Let $\mathcal{U} = \mathcal{H} = L^2(\Omega)$, $J(u) = \|u\|_{L^2(\Omega)}$ and $u^\dagger \in L^2(\Omega)$ with $\|u^\dagger\|_{L^2(\Omega)} = 1$. Then (SSC) reads as $K^*Kw = u^\dagger$ and we see immediately that w does not need to be a multiple of u^\dagger in this case. We now want to look more closely into the connection between (2.18) and (SC).

Proposition 2.8. *Let $A \in \mathcal{L}(\mathcal{U}, \mathcal{H})$ and $f \in \mathcal{H}$, and define $K \in \mathcal{L}(\mathcal{U}, \mathbb{R})$ with $Ku := \langle A^*f, u \rangle$ for all $u \in \mathcal{U}$. Then the singular vector condition (2.18) for K is equivalent to the source condition (SC)*

$$A^*v_\lambda \in \partial J(u_\lambda),$$

for $v_\lambda = J(u_\lambda)\langle A^*f, u_\lambda \rangle f$ and $|\langle A^*f, u_\lambda \rangle| = 1$.

Proof. With the adjoint operator K^* of K being given as

$$K^*c = cA^*f,$$

for all $c \in \mathbb{R}$, the singular vector condition (2.18) reads as

$$\lambda \langle A^*f, u_\lambda \rangle A^*f \in \partial J(u_\lambda),$$

for $|\langle A^*f, u_\lambda \rangle| = 1$. Due to (2.19) we further conclude $\lambda = J(u_\lambda)$, which verifies the proposition. \square

Remark 2.9. Note that we either have $\langle A^*f, u_\lambda \rangle = 1$ or $\langle A^*f, u_\lambda \rangle = -1$, due to the normalisation constraint $|\langle A^*f, u_\lambda \rangle| = 1$. Without loss of generality we assume $\langle A^*f, u_\lambda \rangle = 1$; in this case, the element v_λ in the source condition simplifies to $v_\lambda = J(u_\lambda)f$.

We summarise the assumptions that will be used throughout this work:

Assumption 2.10 (Setup).

- \mathcal{U} is a Banach space.
- \mathcal{H} is a Hilbert space.
- $K \in \mathcal{L}(\mathcal{U}, \mathcal{H})$.
- $J : \text{dom}(J) \subseteq \mathcal{U} \rightarrow \mathbb{R} \cup \{+\infty\}$ is a proper, lower semi-continuous, convex, and absolutely one-homogeneous functional.
- $\ker(K) \cap \ker(J) = \{0\}$.
- $f \in \mathcal{H}$ and $\langle K^*f, v \rangle = 0$ for all $v \in \ker(J)$. Then we can assume $u(0) = u^0 = 0$ for the initial value of the ISS flow.

3. Decomposition of Generalised Singular Vectors

Throughout this section we use Assumption 2.10. In addition we assume that the input data f for the inverse scale space flow (2.10) is represented by a linear combination of generalised singular vectors, i.e. $f = \sum_{j=1}^n \gamma_j K u_{\lambda_j}$ where $\{u_{\lambda_j}\}_{j=1}^n$ is a finite set of generalised singular vectors for J and K and $\gamma_j \in \mathbb{R}$, $j = 1, \dots, n$. The main goal is to investigate if and under which conditions the ISS flow will give a perfect decomposition into the singular vectors representing f , i.e. when do we obtain

$$u(t) = \begin{cases} 0, & 0 \leq t < t_1, \\ \sum_{j=1}^k \gamma_j u_{\lambda_j}, & t_k \leq t < t_{k+1} \text{ for } k = 1, \dots, n-1, \\ \sum_{j=1}^n \gamma_j u_{\lambda_j}, & t_n \leq t, \end{cases} \quad (3.1)$$

for fixed $t_k > 0$, $k = 1, \dots, n$. We will introduce two conditions that will be imposed on the singular vectors in order for the ISS flow to behave this way. In addition we discuss the possibility of singular vector fusion, i.e. the possibility that two or more singular vectors can add up to another singular vector, which would make a decomposition impossible due to Theorem 2.6. Our results will be compared to the achievements of [62] which deals with signal denoising and data represented by two singular vectors.

3.1. Orthogonality Condition

We give a sufficient orthogonality condition on the generalised singular vectors $\{u_{\lambda_j}\}_{j=1}^n$ representing the data $f = \sum_{j=1}^n \gamma_j K u_{\lambda_j}$, for the ISS flow to satisfy the necessary orthogonality condition (2.11).

Let the data for the inverse scale space flow be given as $f = \sum_{j=1}^n \gamma_j K u_{\lambda_j}$ where $\{u_{\lambda_j}\}_{j=1}^n$ is a set of generalised singular vectors. Assume that the inverse scale space flow gives a decomposition into the singular vectors of f with no loss of contrast, i.e. the solution of the ISS flow is given by (3.1). The necessary orthogonality condition (2.11) then states that

$$\begin{aligned} 0 &= \left\langle K^* \left(\sum_{j=1}^n \gamma_j K u_{\lambda_j} - \sum_{j=1}^k \gamma_j K u_{\lambda_j} \right), \sum_{j=1}^k \gamma_j u_{\lambda_j} \right\rangle \\ &= \left\langle \sum_{j=k+1}^n \gamma_j K u_{\lambda_j}, \sum_{j=1}^k \gamma_j K u_{\lambda_j} \right\rangle \end{aligned}$$

for $k = 1, 2, \dots, n-1$. A sufficient condition for this to hold is K -orthogonality between the singular vectors:

$$\langle K u_{\lambda_j}, K u_{\lambda_k} \rangle = 0, \quad j \neq k. \quad (\text{OC})$$

This orthogonality condition will be used throughout the entire paper.

3.2. (SUB0) Condition

K -orthogonality between the singular vectors representing f will not be enough to guarantee that the solution of the ISS flow is given by (3.1). We therefore introduce

an additional condition on the subgradients of the singular vectors that we name the **(SUB0)** condition. It is a condition on the sum of the subgradients of a set of singular vectors which turns out to be a kind of linearity condition for ∂J in the singular vectors. The **(SUB0)** condition furthermore implies linearity of J in the singular vectors.

Definition 3.1 (SUB0 Condition). Let $\{u_{\lambda_j}\}_{j=1}^n$ be a set of singular vectors of J with corresponding singular values $\{\lambda_j\}_{j=1}^n$. We say that $\{u_{\lambda_j}\}_{j=1}^n$ satisfy the **(SUB0)** condition if

$$\sum_{j=1}^k \lambda_j K^* K u_{\lambda_j} \in \partial J(0), \forall k \in \{1; n\}, \quad (\text{SUB0})$$

where $k \in \{1; n\}$ means $k = 1, 2, \dots, n$.

Remark 3.2. In the definition of the **(SUB0)** condition we use the notation $k \in \{1; n\}$ which means $k = 1, 2, \dots, n$. We will use this notation throughout the paper. We want to emphasise that in order for a set of singular vectors to satisfy the **(SUB0)** condition it is necessary that the inclusion is satisfied for all $k \in \{1; n\}$. It is not enough that the full sum over the subgradients is included in the subdifferential at zero. We will call the inclusions for $k \in \{1; n-1\}$ the *partial sum conditions* of the **(SUB0)** condition.

The **(SUB0)** condition is equivalent to a linearity condition for the subdifferential for K -orthogonal singular vectors:

Proposition 3.3. Let $\{u_{\lambda_j}\}_{j=1}^n$ be a set of K -normalised singular vectors of J with corresponding singular values $\{\lambda_j\}_{j=1}^n$. Assume that they satisfy the orthogonality condition **(OC)**. Then the **(SUB0)** condition is satisfied if and only if

$$\sum_{j=1}^k \lambda_j K^* K u_{\lambda_j} \in \partial J \left(\sum_{j=1}^k c_j u_{\lambda_j} \right), \quad c_j \geq 0, \forall k \in \{1; n\}.$$

Proof. Proof of " \Leftarrow ": This is straightforward using the subdifferential characterisation (2.6).

Proof of " \Rightarrow ": Using the **(SUB0)** condition, the triangle inequality (2.4) for J and the orthogonality condition **(OC)** we get

$$\begin{aligned} J(v) - J \left(\sum_{j=1}^k c_j u_{\lambda_j} \right) &= \left\langle \sum_{j=1}^k \lambda_j K^* K u_{\lambda_j}, v - \sum_{j=1}^k c_j u_{\lambda_j} \right\rangle \\ &\geq J(v) - \sum_{j=1}^k c_j J(u_{\lambda_j}) - J(v) + \sum_{j=1}^k c_j J(u_{\lambda_j}) \\ &= 0 \end{aligned}$$

for $c_j \geq 0$, $j = 1, \dots, k$, for all $v \in \text{dom}(J)$. This shows that $\sum_{j=1}^k \lambda_j K^* K u_{\lambda_j} \in \partial J \left(\sum_{j=1}^k c_j u_{\lambda_j} \right)$ for all $k \in \{1; n\}$. \square

We would like to make a comparison to the (LIS) condition in [62, Definition 7]. LIS stands for linearity in the subdifferential. The (LIS) condition is in principle the same as the (SUB0) condition through Proposition 3.3, except that the (LIS) condition is only defined for a pair of two singular vectors and additionally requires the linearity in the subdifferential to hold for negative coefficients $c_j < 0$ as well. (LIS) turns out to imply directly orthogonality between the singular vectors, see [62, Proposition 4].

The (SUB0) condition furthermore implies some kind of linearity of J for K -orthogonal singular vectors:

Proposition 3.4. *Let $\{u_{\lambda_j}\}_{j=1}^n$ be a set of K -normalised singular vectors of J with corresponding singular values $\{\lambda_j\}_{j=1}^n$. Assume that they satisfy the orthogonality condition (OC) and the (SUB0) condition. Then*

$$J\left(\sum_{j=1}^n c_j u_{\lambda_j}\right) = \sum_{j=1}^n c_j J(u_{\lambda_j})$$

for $c_j \geq 0$, $j = 1, \dots, n$.

Proof. Using Proposition 3.3, the properties of the subdifferential for absolutely one-homogeneous functionals, and (OC) we get

$$J\left(\sum_{j=1}^n c_j u_{\lambda_j}\right) = \left\langle \sum_{j=1}^n \lambda_j K^* K u_{\lambda_j}, \sum_{j=1}^n c_j u_{\lambda_j} \right\rangle = \sum_{j=1}^n c_j J(u_{\lambda_j}).$$

□

Remark 3.5. Note that the linearity in J is only satisfied for non-negative coefficients c_j and the proposition actually holds even if (SUB0) is only satisfied for $k = n$.

Remark 3.6. Linearity of J does not imply the (SUB0) condition under (OC), see Example 3.8 for a counter example.

In order to give an intuitive idea of what the orthogonality condition (OC) and the (SUB0) condition require of the singular vectors, we give three examples:

Example 3.7. Let us again consider $K = I : \ell^1(\mathbb{R}^m) \rightarrow \ell^2(\mathbb{R}^m)$ and $J(u) = \|Wu\|_{\ell^1(\mathbb{R}^m)}$ as in Example 2.4. In order for two singular vectors u_{λ_1} and u_{λ_2} to be orthogonal we need

$$0 = \langle u_{\lambda_1}, u_{\lambda_2} \rangle = \langle u_{\lambda_1}, W^* W u_{\lambda_2} \rangle = \langle W u_{\lambda_1}, W u_{\lambda_2} \rangle.$$

Hence we can just check that $W u_{\lambda_1}$ and $W u_{\lambda_2}$ are orthogonal in $\ell^2(\mathbb{R}^m)$. For the (SUB0) condition we have

$$\begin{aligned} \lambda_1 u_{\lambda_1} + \lambda_2 u_{\lambda_2} &= \lambda_1 W^* W u_{\lambda_1} + \lambda_2 W^* W u_{\lambda_2} \\ &= W^*(\text{sign}(W u_{\lambda_1}) + \text{sign}(W u_{\lambda_2})). \end{aligned}$$

If $W u_{\lambda_1}$ and $W u_{\lambda_2}$ have non-overlapping supports then

$$\lambda_1 u_{\lambda_1} + \lambda_2 u_{\lambda_2} = W^*(\text{sign}(W u_{\lambda_1} + W u_{\lambda_2})) \in \partial J(0)$$

and the (SUB0) condition is satisfied. Note that in this case also the orthogonality condition (OC) is satisfied.

Example 3.8. Let $J(u) = \|u\|_{\ell^1(\mathbb{R}^m)}$ with $\partial J(u) = \text{sign}(u)$, see Example 2.4. We look at a convolution problem where $K : \ell^1(\mathbb{R}^m) \supseteq \ell^2(\mathbb{R}^m) \rightarrow \ell^2(\mathbb{R}^m)$ is the convolution operator defined by

$$(Ku)_k = \sum_{j=-1}^1 u_{k-j} g_j, \quad g = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.$$

The adjoint operator K^* is determined by

$$(K^*v)_j = \sum_{k=-1}^1 v_{j+k} g_k.$$

Two examples of K -normalised singular vectors are

$$u_{\lambda_1} = (0, 0, 1, 0, 0)^T \quad \text{and} \quad u_{\lambda_2} = \frac{1}{\sqrt{2}}(0, 1, 0, -1, 0)^T.$$

We have

$$\begin{aligned} Ku_{\lambda_1} &= (0, 1/\sqrt{2}, 1/\sqrt{2}, 0, 0)^T, & \|Ku_{\lambda_1}\|_{\ell^1(\mathbb{R}^m)} &= 1, \\ Ku_{\lambda_2} &= (1/2, 1/2, -1/2, -1/2, 0)^T, & \|Ku_{\lambda_2}\|_{\ell^1(\mathbb{R}^m)} &= 1, \end{aligned}$$

and

$$\begin{aligned} J(u_{\lambda_1})K^*Ku_{\lambda_1} &= (0, 1/2, 1, 1/2, 0)^T \in \text{sign}(u_{\lambda_1}), \\ J(u_{\lambda_2})K^*Ku_{\lambda_2} &= (1/2, 1, 0, -1, -1/2)^T \in \text{sign}(u_{\lambda_2}). \end{aligned}$$

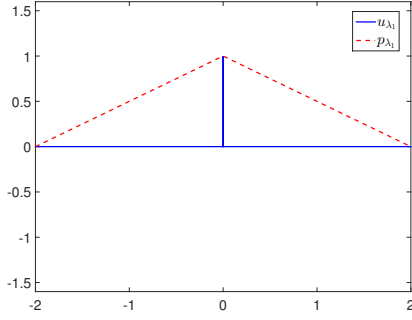
The two singular vectors are K -orthogonal but they do not satisfy the **(SUB0)** condition since

$$J(u_{\lambda_1})K^*Ku_{\lambda_1} + J(u_{\lambda_2})K^*Ku_{\lambda_2} = (1/2, \textcolor{red}{3/2}, 1, -1/2, -1/2)^T \notin \partial J(0) = [-1, 1].$$

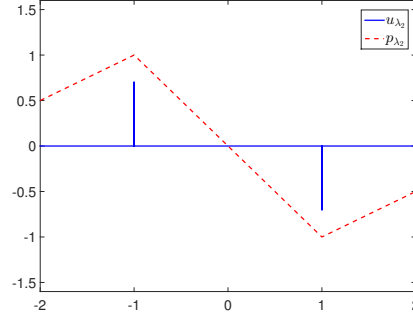
This example also shows that Proposition 3.4 only holds one way since for $c_1, c_2 \geq 0$ we have

$$\begin{aligned} J(c_1 u_{\lambda_1} + c_2 u_{\lambda_2}) &= \left\| (0, c_2/\sqrt{2}, c_1, -c_2/\sqrt{2}, 0)^T \right\|_{\ell^1(\mathbb{R}^m)} = c_1 + 2c_2/\sqrt{2} \\ &= c_1 \cdot 1 + c_2 \cdot 2/\sqrt{2} = c_1 J(u_{\lambda_1}) + c_2 J(u_{\lambda_2}). \end{aligned}$$

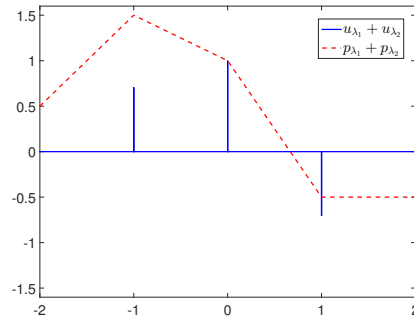
Hence the linearity of J in the singular vectors does not imply that the **(SUB0)** condition is satisfied. In Figure 1 we see the two singular vectors and their corresponding subgradients. We see that the subgradients contain values between -1 and 1 and are exactly -1 at the negative peaks and 1 at the positive peaks. In Figure 1c we see the sum of the singular vectors that contains three peaks. The sum of the subgradients contains values above one and is therefore not in the subdifferential of J at zero. The problem is that the two positive peaks of the singular vectors get too close, which makes the sum of the subgradients exceed one.



(a) First singular vector and subgradient.



(b) Second singular vector and subgradient.



(c) Sum of singular vectors and subgradients.

Figure 1: *Convolution example.* In Figures 1a and 1b we see two K -orthogonal singular vectors. Their subgradients reach 1, respectively -1 , for positive, respectively negative, peaks. Between the peaks the value is in $[-1, 1]$. In Figure 1c we see the sum of the two singular vectors and the sum of the subgradients. The **(SUB0)** condition is not satisfied since the sum of the subgradients exceeds one. The problem is the two positive peaks of the singular vectors getting too close.

Example 3.9. We consider $K = I : BV([0, 1]) \rightarrow L^2([0, 1])$, $J(u) = \text{TV}_*(u)$, and the two Haar wavelets u_{λ_1} and u_{λ_2} as defined in Example 2.3. By adding these two singular vectors we get

$$(u_{\lambda_1} + u_{\lambda_2})(x) = \begin{cases} 1 + \sqrt{2}, & 0 \leq x < 1/4, \\ 1 - \sqrt{2}, & 1/4 \leq x < 1/2, \\ -1, & 1/2 \leq x \leq 1, \\ 0, & \text{otherwise.} \end{cases}$$

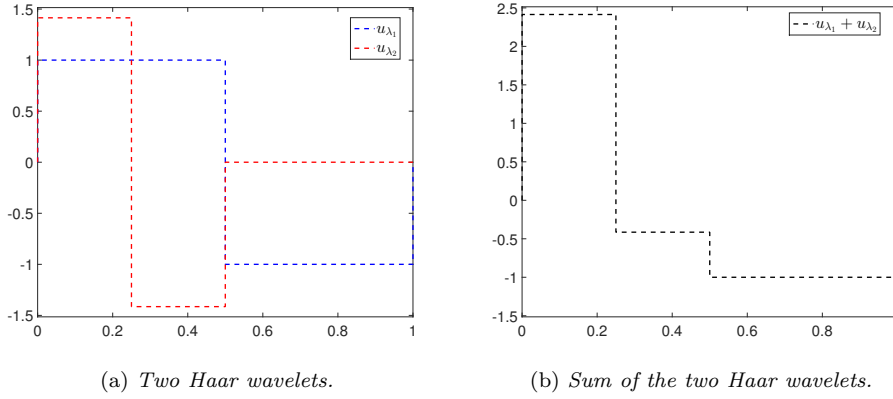


Figure 2: TV_* -denoising example. In Figure 2a we see two Haar wavelets which are orthogonal singular vectors for the TV_* -functional and the embedding operator $I : BV([0, 1]) \rightarrow L^2([0, 1])$. The singular vectors do not satisfy the **(SUB0)** condition by Proposition 3.4, since $TV_*(u_{\lambda_1} + u_{\lambda_2}) < TV_*(u_{\lambda_1}) + TV_*(u_{\lambda_2})$.

The total variation of the sum is then

$$\begin{aligned}
 TV_*(u_{\lambda_1} + u_{\lambda_2}) &= \sup_{\substack{\varphi \in C^\infty([0, 1]), \\ \|\varphi\|_{L^\infty([0, 1])} \leq 1}} \int_0^1 (u_{\lambda_1} + u_{\lambda_2})(x) \varphi'(x) dx, \\
 &= \sup_{\substack{\varphi \in C^\infty([0, 1]), \\ \|\varphi\|_{L^\infty([0, 1])} \leq 1}} \left\{ (1 + \sqrt{2}) \int_0^{1/4} \varphi'(x) dx + (1 - \sqrt{2}) \int_{1/4}^{1/2} \varphi'(x) dx \right. \\
 &\quad \left. - \int_{1/2}^1 \varphi'(x) dx \right\} \\
 &= \sup_{\substack{\varphi \in C^\infty([0, 1]), \\ \|\varphi\|_{L^\infty([0, 1])} \leq 1}} \left\{ -(1 + \sqrt{2})\varphi(0) + 2\sqrt{2}\varphi(1/4) + (2 - \sqrt{2})\varphi(1/2) \right. \\
 &\quad \left. - \varphi(1) \right\} \\
 &\leq 1 + \sqrt{2} + 2\sqrt{2} + 2 - \sqrt{2} + 1 = 4 + 2\sqrt{2} \\
 &= TV_*(u_{\lambda_1}) + \frac{1}{2}TV_*(u_{\lambda_2}) \\
 &< TV_*(u_{\lambda_1}) + TV_*(u_{\lambda_2}).
 \end{aligned}$$

By Proposition 3.4, u_{λ_1} and u_{λ_2} do not satisfy the **(SUB0)** condition. In Figure 2a we see the two Haar wavelets plotted. In Figure 2b we see the sum of the two Haar wavelets. At $x = 1/2$ we have a second jump downwards. This is at the place where the two Haar wavelets have overlapping jumps in opposite directions. That is exactly what makes the TV_* -functional non-linear in the singular vectors.

In the following we are going to provide an alternative *dual norm formulation* of the

(SUB0) condition under a norm-inequality assumption on J .

Definition 3.10 (Norm-inequality Assumption). J satisfies the *norm-inequality assumption* if there exists a positive constant $c_0 > 0$ such that

$$c_0 \|u\|_{\mathcal{U}} \leq J(u), \quad \forall u \in \ker(J)^\perp. \quad (3.2)$$

Under the norm-inequality assumption we are able to show a relation between the subdifferential at zero and the dual norm:

Proposition 3.11. *Let J be absolutely one-homogeneous and satisfy the norm-inequality assumption (3.2). Then*

$$\|\omega^*\|_{\mathcal{U}^*} \leq c_0 \text{ and } \langle \omega^*, v \rangle = 0, \forall v \in \ker(J) \Rightarrow \omega^* \in \partial J(0).$$

Proof. Let $\|\omega^*\|_{\mathcal{U}^*} \leq c_0$ and $\langle \omega^*, v \rangle = 0$ for all $v \in \ker(J)$. We need to show $\langle \omega^*, \omega \rangle \leq J(\omega)$ for all $\omega \in \mathcal{U}$. We can already exclude all $\omega \in \ker(J)$ since in this case we have $\langle \omega^*, \omega \rangle = 0$ by assumption and $J(\omega) = 0$. For $\omega \in \ker(J)^\perp$ we get

$$\langle \omega^*, \omega \rangle \leq \|\omega^*\|_{\mathcal{U}^*} \|\omega\|_{\mathcal{U}} \leq c_0 \|\omega\|_{\mathcal{U}} \leq J(\omega)$$

by (3.2). Hence for $\omega = \omega_1 + \omega_2$ with $\omega_1 \in \ker(J)$ and $\omega_2 \in \ker(J)^\perp$ we have

$$\langle \omega^*, \omega \rangle = \langle \omega^*, \omega_2 \rangle \leq J(\omega_2) = J(\omega),$$

where the last equality follows from (2.5). \square

Now we see that if

$$\left\| \sum_{j=1}^k \lambda_j K^* K u_{\lambda_j} \right\|_{\mathcal{U}^*} \leq c_0 \quad (3.3)$$

for $k \in \{1; n\}$ then the (SUB0) condition is satisfied, since

$$\left\langle \sum_{j=1}^k \lambda_j K^* K u_{\lambda_j}, v \right\rangle = 0, \quad \forall v \in \ker(J)$$

is already satisfied by (2.8) and $\lambda_j K^* K u_{\lambda_j} \in \partial J(u_{\lambda_j})$ for $j = 1, \dots, k$. Hence we can use (3.3) as an alternative to the (SUB0) condition for numerical implementations when the norm-inequality assumption is satisfied.

Remark 3.12. Note that if $J(u) = \|u\|_{\mathcal{U}}$ then

$$\|\omega^*\|_{\mathcal{U}^*} \leq 1 \text{ and } \langle \omega^*, v \rangle = 0, \forall v \in \ker(J) \Rightarrow \omega^* \in \partial J(0)$$

by Proposition 3.11 with $c_0 = 1$ and for $\omega^* \in \partial J(0)$ we have

$$\|\omega^*\|_{\mathcal{U}^*} = \sup_{\|v\|_{\mathcal{U}} \leq 1} \langle \omega^*, v \rangle \leq \sup_{\|v\|_{\mathcal{U}} \leq 1} J(v) = \sup_{J(v) \leq 1} J(v) = 1.$$

Hence in this case the (SUB0) condition is equivalent to

$$\left\| \sum_{j=1}^n \lambda_j K^* K u_{\lambda_j} \right\|_{\mathcal{U}^*} \leq 1.$$

Example 3.13. With this final **(SUB0)**-example we want to highlight that given the orthogonality condition **(OC)**, the one-homogeneous analogue to Showalter's method, i.e. the ISS flow (2.10) with $J(u) = \|u\|_{L^2(\Omega)}$, $\mathcal{U} = L^2(\Omega)$, satisfies the **(SUB0)** condition only for $k = 1$. We see this by considering **(SUB0)** in this particular case. We have

$$\left\| \sum_{j=1}^k \lambda_j K^* K u_{\lambda_j} \right\|_{L^2(\Omega)} = \left\| \sum_{j=1}^k \frac{u_{\lambda_j}}{\|u_{\lambda_j}\|_{L^2(\Omega)}} \right\|_{L^2(\Omega)},$$

due to $\partial\|u\|_{L^2(\Omega)} = u/\|u\|_{L^2(\Omega)}$ for $\|u\|_{L^2(\Omega)} \neq 0$. Due to the K -orthogonality in combination with the singular vector condition (2.18) we observe $0 = \lambda_i \langle K u_{\lambda_i}, K u_{\lambda_j} \rangle = \langle u_{\lambda_i}, u_{\lambda_j} \rangle / \|u_{\lambda_i}\|_{L^2(\Omega)}$, which implies $\langle u_{\lambda_i}, u_{\lambda_j} \rangle = 0$. Hence, we have

$$\begin{aligned} \left\| \sum_{j=1}^k \frac{u_{\lambda_j}}{\|u_{\lambda_j}\|_{L^2(\Omega)}} \right\|_{L^2(\Omega)}^2 &= \left\langle \sum_{i=1}^k \frac{u_{\lambda_i}}{\|u_{\lambda_i}\|_{L^2(\Omega)}}, \sum_{j=1}^k \frac{u_{\lambda_j}}{\|u_{\lambda_j}\|_{L^2(\Omega)}} \right\rangle \\ &= \sum_{i=1}^k \frac{\langle u_{\lambda_i}, u_{\lambda_i} \rangle}{\|u_{\lambda_i}\|_{L^2(\Omega)} \|u_{\lambda_i}\|_{L^2(\Omega)}} = k, \end{aligned}$$

and therefore

$$\left\| \sum_{j=1}^k \lambda_j K^* K u_{\lambda_j} \right\|_{L^2(\Omega)} = \sqrt{k} \geq 1,$$

for $k \in \mathbb{N}$. Thus, according to Remark 3.12, the **(SUB0)** condition is only satisfied for $k = 1$.

3.3. Inverse Scale Space Decomposition

We now provide our main result which states that the ISS flow gives a perfect decomposition into the generalised singular vectors representing the data of the flow given that **(OC)** and **(SUB0)** are satisfied.

Theorem 3.14 (Inverse Scale Space Decomposition). *Let $\{u_{\lambda_j}\}_{j=1}^n$ be a set of K -normalised singular vectors of J with corresponding singular values $\{\lambda_j\}_{j=1}^n$. Assume that they satisfy the orthogonality condition **(OC)** and the **(SUB0)** condition. Then, if the data f is given by $f = \sum_{j=1}^n \gamma_j K u_{\lambda_j}$ for positive constants $\{\gamma_j\}_{j=1}^n$, a solution of the inverse scale space flow (2.10) is given by*

$$u(t) = \begin{cases} 0, & 0 \leq t < t_1, \\ \sum_{j=1}^k \gamma_j u_{\lambda_j}, & t_k \leq t < t_{k+1} \text{ for } k = 1, \dots, n-1, \\ \sum_{j=1}^n \gamma_j u_{\lambda_j}, & t_n \leq t, \end{cases}$$

where $t_k = \lambda_k / \gamma_k$ and $t_k < t_{k+1}$.

Proof. The proof is divided into the following three steps corresponding to the three different cases for $u(t)$:

Step 1: $0 \leq t < t_1$.

Step 2: $t_k \leq t < t_{k+1}$ for $k = 1, \dots, n-1$

Step 3: $t_n \leq t$.

For every step we have to show that the solution satisfies the inverse scale space flow.

Step 1: Let $0 \leq t < t_1$. Then $u(t) = 0$, and from the inverse scale space flow we get

$$\partial_t p(t) = K^* f, \quad p(0) = 0.$$

Integrating $\int_0^t \partial_\tau p(\tau) d\tau$ yields

$$p(t) = tK^* f = t \left(\sum_{j=1}^n \frac{\gamma_j}{\lambda_j} p_{\lambda_j} \right),$$

where we have used the notation $p_{\lambda_j} = \lambda_j K^* K u_{\lambda_j}$. We need to show that $p(t) \in \partial J(u(t)) = \partial J(0)$. Using the characterisation (2.6) of the subdifferential we need to show that $\langle p(t), v \rangle \leq J(v)$ for all $v \in \mathcal{U}$. We have

$$\langle p(t), v \rangle = t \left\langle \sum_{j=1}^n \frac{\gamma_j}{\lambda_j} p_{\lambda_j}, v \right\rangle.$$

Hence we need to show

$$t \left\langle \sum_{j=1}^n \frac{\gamma_j}{\lambda_j} p_{\lambda_j}, v \right\rangle \leq J(v), \quad \forall v \in \mathcal{U}. \quad (3.4)$$

If $\left\langle \sum_{j=1}^n \frac{\gamma_j}{\lambda_j} p_{\lambda_j}, v \right\rangle \leq 0$ then inequality (3.4) is trivially satisfied since $J(v) \geq 0$ for all $v \in \mathcal{U}$. If $\left\langle \sum_{j=1}^n \frac{\gamma_j}{\lambda_j} p_{\lambda_j}, v \right\rangle > 0$ then

$$t \left\langle \sum_{j=1}^n \frac{\gamma_j}{\lambda_j} p_{\lambda_j}, v \right\rangle < t_1 \left\langle \sum_{j=1}^n \frac{\gamma_j}{\lambda_j} p_{\lambda_j}, v \right\rangle = \langle p_{\lambda_1}, v \rangle + t_1 \left\langle \sum_{j=2}^n \frac{\gamma_j}{\lambda_j} p_{\lambda_j}, v \right\rangle$$

If $\left\langle \sum_{j=2}^n \frac{\gamma_j}{\lambda_j} p_{\lambda_j}, v \right\rangle \leq 0$ inequality (3.4) is satisfied since $p_{\lambda_1} \in \partial J(0)$. If $\left\langle \sum_{j=2}^n \frac{\gamma_j}{\lambda_j} p_{\lambda_j}, v \right\rangle > 0$ then

$$\begin{aligned} \langle p_{\lambda_1}, v \rangle + t_1 \left\langle \sum_{j=2}^n \frac{\gamma_j}{\lambda_j} p_{\lambda_j}, v \right\rangle &< \langle p_{\lambda_1}, v \rangle + t_2 \left\langle \sum_{j=2}^n \frac{\gamma_j}{\lambda_j} p_{\lambda_j}, v \right\rangle \\ &= \langle p_{\lambda_1} + p_{\lambda_2}, v \rangle + t_2 \left\langle \sum_{j=3}^n \frac{\gamma_j}{\lambda_j} p_{\lambda_j}, v \right\rangle \end{aligned}$$

If $\left\langle \sum_{j=3}^n \frac{\gamma_j}{\lambda_j} p_{\lambda_j}, v \right\rangle \leq 0$ then inequality (3.4) is satisfied since $p_{\lambda_1} + p_{\lambda_2} \in \partial J(0)$ by the **(SUB0)** condition. If $\left\langle \sum_{j=3}^n \frac{\gamma_j}{\lambda_j} p_{\lambda_j}, v \right\rangle > 0$ we continue the process. The process

terminates when we reach

$$\left\langle \sum_{j=1}^n p_{\lambda_j}, v \right\rangle,$$

which is less than or equal to $J(v)$ since $\sum_{j=1}^n p_{\lambda_j} \in \partial J(0)$ by the **(SUB0)** condition. In any case we have shown that $p(t) \in \partial J(0)$ for $0 \leq t < t_1$.

Step 2: Let $t_k \leq t < t_{k+1}$. Then $u(t) = \sum_{j=1}^k \gamma_j u_{\lambda_j}$ and from the inverse scale space flow we get

$$\partial_t p(t) = K^* \left(f - \sum_{j=1}^k \gamma_j K u_{\lambda_j} \right) = \sum_{j=k+1}^n \gamma_j K^* K u_{\lambda_j} = \sum_{j=k+1}^n \frac{\gamma_j}{\lambda_j} p_{\lambda_j}.$$

A continuous extension of p at $t = t_k$ is given by

$$p(t) = \sum_{j=1}^k p_{\lambda_j} + t \left(\sum_{j=k+1}^n \frac{\gamma_j}{\lambda_j} p_{\lambda_j} \right).$$

We require $p(t) \in \partial J(u(t)) = \partial J \left(\sum_{j=1}^k \gamma_j u_{\lambda_j} \right)$. Using the characterisation (2.6) of the subdifferential we need to show

$$\left\langle p(t), \sum_{j=1}^k \gamma_j u_{\lambda_j} \right\rangle = J \left(\sum_{j=1}^k \gamma_j u_{\lambda_j} \right) \quad \text{and} \quad \langle p(t), v \rangle \leq J(v), \quad \forall v \in \mathcal{U}. \quad (3.5)$$

For the first criterion of (3.5) we get

$$\begin{aligned} \left\langle p(t), \sum_{j=1}^k \gamma_j u_{\lambda_j} \right\rangle &= \left\langle \sum_{j=1}^k p_{\lambda_j}, \sum_{j=1}^k \gamma_j u_{\lambda_j} \right\rangle + t \left\langle \sum_{j=k+1}^n \frac{\gamma_j}{\lambda_j} p_{\lambda_j}, \sum_{j=1}^k \gamma_j u_{\lambda_j} \right\rangle \\ &= \sum_{j=1}^k \gamma_j J(u_{\lambda_j}) = J \left(\sum_{j=1}^k \gamma_j u_{\lambda_j} \right). \end{aligned}$$

We have used the formula for $p(t)$, the properties of singular vectors, the orthogonality condition **(OC)** and for the last equality Proposition 3.4. For the second criterion of (3.5) we can use the same process as for **Step 1**.

Step 3: Let $t_n \leq t$. Then $u(t) = \sum_{j=1}^n \gamma_j u_{\lambda_j}$ and from the inverse scale space flow we get

$$\partial_t p(t) = K^* \left(f - \sum_{j=1}^n \gamma_j K u_{\lambda_j} \right) = 0.$$

A continuous extension of p at $t = t_n$ is

$$p(t) = \sum_{j=1}^n p_{\lambda_j}.$$

We need to show that $p(t) \in \partial J(u(t)) = \partial J \left(\sum_{j=1}^n \gamma_j u_{\lambda_j} \right)$. This follows directly from Proposition 3.3. \square

Remark 3.15. Note that for $n = 2$ the **(SUB0)** condition is both necessary and sufficient under the orthogonality condition **(OC)**. In this case we can compute

$$p(t) = \begin{cases} tK^*f, & 0 \leq t < t_1, \\ p_{\lambda_1} + t\gamma_2 K^* K u_{\lambda_2}, & t_1 \leq t < t_2, \\ p_{\lambda_1} + p_{\lambda_2}, & t_2 \leq t. \end{cases}$$

Hence we need $p(t) \in \partial J(u(t))$ for $t_2 \leq t$, i.e.

$$p_{\lambda_1} + p_{\lambda_2} \in \partial J(\gamma_1 u_{\lambda_1} + \gamma_2 u_{\lambda_2}) \subset \partial J(0).$$

This shows that the **(SUB0)** condition is necessary. In fact, for arbitrary n we always need $\sum_{j=1}^n p_{\lambda_j} \in \partial J(0)$ for the decomposition to happen. This is, however, not enough, as we will see in Example 3.18.

Remark 3.16. If u_{λ_j} is a singular vector then also $-u_{\lambda_j}$ is a singular vector. Hence if $\gamma_j < 0$, we can just pull the negative sign of γ_j onto u_{λ_j} and regard this as a singular vector with a positive contribution to f . Then Theorem 3.14 still holds true with $t_k = \lambda_k/|\gamma_k|$ and the **(SUB0)** condition being replaced by

$$\sum_{j=1}^k \frac{\gamma_j}{|\gamma_j|} \lambda_j K^* K u_{\lambda_j} \in \partial J(0), \quad \forall k \in \{1; n\}.$$

Remark 3.17. In Example 3.13 we have shown that the **(SUB0)** condition is never satisfied for $n > 1$ for $J(u) = \|u\|_{L^2(\Omega)}$, $\mathcal{U} = L^2(\Omega)$. Since by Remark 3.15 the full sum over the subgradients of the singular vectors always has to be in the subdifferential of J at zero in order for the ISS flow to give a decomposition into the K -orthogonal singular vectors representing f , we then conclude that the inverse scale space decomposition can never happen in this case without re-weighting its coefficients.

We give a comparison to a decomposition result in [62, Theorem 2] under the earlier mentioned **(LIS)** condition [62, Definition 7] and orthogonality condition **(FO)** [62, Definition 5]. As previously mentioned the **(LIS)** condition requires linearity in the subdifferential for both positive and negative coefficients. The orthogonality condition is formulated a bit differently but boils down to the same as **(OC)** and is actually automatically satisfied under the **(LIS)** condition when J is absolutely one-homogeneous. It is then shown that the (forward) scale space flow

$$\partial_t u(t) = -p(t), \quad p(t) \in \partial J(u(t)), \quad u(0) = f$$

gives a decomposition into singular vectors under the **(LIS)** and **(FO)** conditions when f is given as the sum of two singular vectors and J is one-homogeneous. The results are stated for the forward scale space flow and for denoising problems where K is the identity and involve decomposition into two singular vectors only.

To emphasise that the partial sum conditions of the **(SUB0)** condition are necessary we look at the following example:

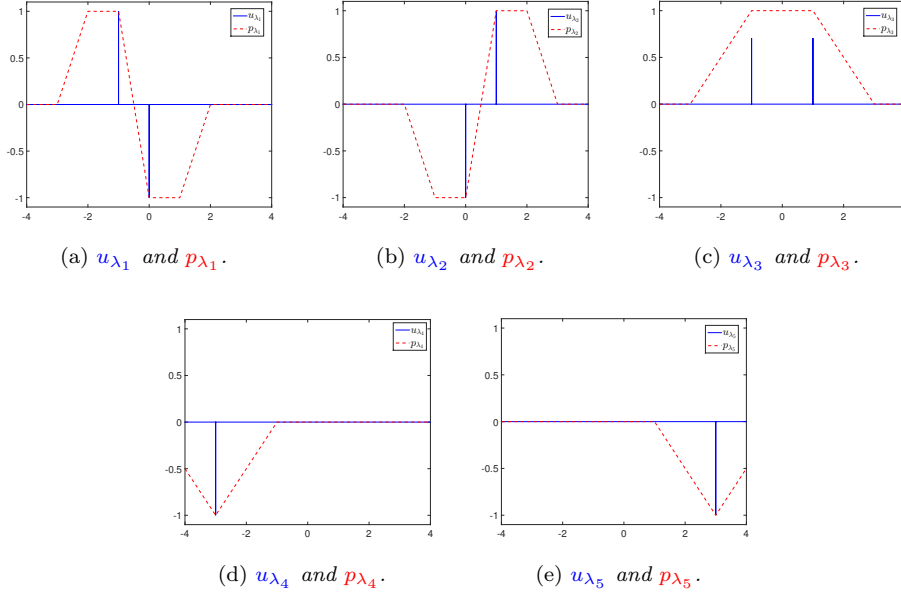


Figure 3: Singular vectors and their corresponding subgradients for Example 3.18

Example 3.18. Let $J(u) = \|u\|_{\ell^1(\mathbb{R}^m)}$ with $\partial J(u) = \text{sign}(u)$, see example 2.4. We reuse the convolution operator $K : \ell^1(\mathbb{R}^m) \rightarrow \ell^2(\mathbb{R}^m)$ from Example 3.8. Define the vectors

$$\begin{aligned} u_{\lambda_1} &= (0, 0, 0, 1, -1, 0, 0, 0)^T, & u_{\lambda_2} &= (0, 0, 0, 0, -1, 1, 0, 0)^T, \\ u_{\lambda_3} &= \frac{1}{\sqrt{2}}(0, 0, 0, 1, 0, 1, 0, 0)^T, & u_{\lambda_4} &= (0, -1, 0, 0, 0, 0, 0, 0)^T, \\ u_{\lambda_5} &= (0, 0, 0, 0, 0, 0, 0, -1)^T. \end{aligned}$$

Using the notation $p_{\lambda_j} = \lambda_j K^* K u_{\lambda_j}$ we get

$$\begin{aligned} p_{\lambda_1} &= (0, 0, 1, 1, -1, -1, 0, 0)^T \in \partial J(u_{\lambda_1}), \\ p_{\lambda_2} &= (0, 0, 0, -1, -1, 1, 1, 0)^T \in \partial J(u_{\lambda_2}), \\ p_{\lambda_3} &= (0, 0, 1/2, 1, 1, 1/2, 0, 0)^T \in \partial J(u_{\lambda_3}), \\ p_{\lambda_4} &= (-1/2, -1, -1/2, 0, 0, 0, 0, 0)^T \in \partial J(u_{\lambda_4}), \\ p_{\lambda_5} &= (0, 0, 0, 0, 0, 0, -1/2, -1)^T \in \partial J(u_{\lambda_5}). \end{aligned}$$

Hence $\{u_{\lambda_j}\}_{j=1}^5$ is a set of five singular vectors. In Figure 3 we see all the singular vectors and their corresponding subgradients. We observe that two peaks of the same sign cannot get too close if we want to obtain a singular vector since the convolution kernel spreads out the peaks. We also see that the subgradients have their maximum magnitude 1 at the peaks.

We note that $\langle K u_{\lambda_i}, K u_{\lambda_j} \rangle = \frac{1}{\lambda_i} \langle p_{\lambda_i}, u_{\lambda_j} \rangle = 0$ for $i \neq j$. Hence the orthogonality condition **(OC)** is satisfied. The singular vectors are also normalised such that

$\|Ku_{\lambda_j}\|_{\ell^2(\mathbb{R}^m)} = 1$ for $j = 1, \dots, 5$ holds true. Now let $\{\gamma_j\}_{j=1}^5$ be such that $\lambda_k/\gamma_k < \lambda_{k+1}/\gamma_{k+1}$ for $k = 1, \dots, 4$. Then if the **(SUB0)** condition is satisfied, Theorem 3.14 tells us that the ISS flow gives a decomposition of the data $f = \sum_{j=1}^5 \gamma_j Ku_{\lambda_j}$ into the singular vectors. But the **(SUB0)** condition is not satisfied since the partial sums over the subgradients are not in the subdifferential of J at zero:

$$\begin{aligned} p_{\lambda_1} + p_{\lambda_2} &= (0, 0, 1, 0, -2, 0, 1, 0, 0)^T \notin \partial J(0), \\ p_{\lambda_1} + p_{\lambda_2} + p_{\lambda_3} &= (0, 0, 3/2, 1, -1, 1, 3/2, 0, 0)^T \notin \partial J(0), \\ p_{\lambda_1} + p_{\lambda_2} + p_{\lambda_3} + p_{\lambda_4} &= (-1/2, -1, 1, 1, -1, 1, 3/2, 0, 0)^T \notin \partial J(0), \\ p_{\lambda_1} + p_{\lambda_2} + p_{\lambda_3} + p_{\lambda_4} + p_{\lambda_5} &= (-1/2, -1, 1, 1, -1, 1, 1, -1, -1/2)^T \in \partial J(0). \end{aligned}$$

Only the full sum over all the subgradients is in the subdifferential of J at zero. In Figure 4 we see some of the partial sums over the subgradients. In Figure 4a we see the sum over the first three subgradients. We see that the sum does not have its maximum magnitude at the peaks and it exceeds the limit of a magnitude of one. Hence the sum of the first three subgradients is not in the subdifferential of J at zero. In Figure 4b we see the sum over all five subgradients. This sum has its maximum magnitude at the peaks and does not exceed the magnitude limit of one. Hence the full sum is in the subdifferential of J at zero.

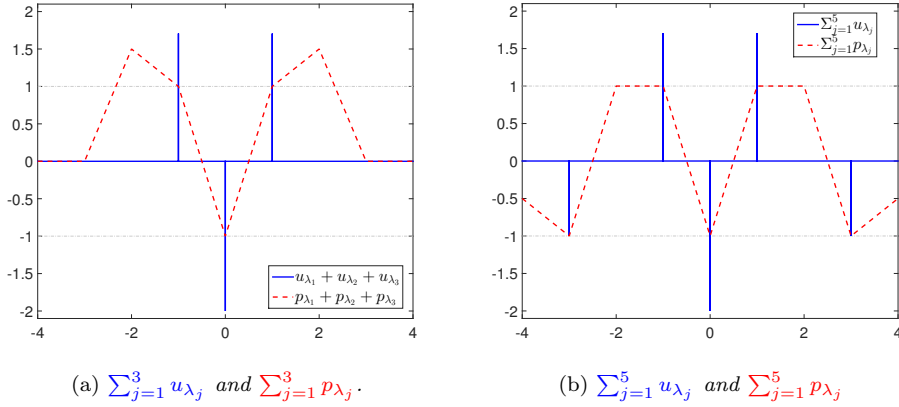


Figure 4: Sum of subgradients for Example 3.18. We see that the partial sum conditions of the **(SUB0)** condition are not satisfied. The sum over the first three subgradients seen in Figure 4a exceeds one and can therefore not be in the subdifferential of J at zero. In Figure 4b the full sum over the subgradients is seen and this is in the subdifferential of J at zero.

We are going to see in Section 5.2 that the first peak reconstructed by the ISS flow in this case is at the middle position. This peak does not match any of the singular vectors u_{λ_j} , $j = 1, \dots, 5$, and we will not obtain the decomposition.

3.4. Singular Vector Fusion

One might think that if two or more singular vectors representing the data f add up to another singular vector, then the ISS flow is not able to give a complete decomposi-

tion into the singular vectors. However, it turns out that under the **(SUB0)** condition and the orthogonality condition **(OC)**, two or more singular vectors cannot add up to another singular vector.

Proposition 3.19 (Impossible Singular Vector Fusion). *Let $\{u_{\lambda_j}\}_{j=1}^n$ be a set of K -normalised singular vectors of J with corresponding singular values $\{\lambda_j\}_{j=1}^n$. Assume that they satisfy the orthogonality condition **(OC)** and the **(SUB0)** condition. Let the data f be given by $f = \sum_{j=1}^n \gamma_j K u_{\lambda_j}$, where $\gamma_j > 0$ and $\lambda_j/\gamma_j < \lambda_{j+1}/\gamma_{j+1}$. Let $\{u_{\lambda_{j_k}}\}_{k=1}^m$, $m \in \{2; n\}$, be a subset of the singular vectors ordered so that $\lambda_{j_k}/\gamma_{j_k} < \lambda_{j_{k+1}}/\gamma_{j_{k+1}}$. Then no such subset of singular vectors can add up to another singular vector, i.e.*

$$\frac{J\left(\sum_{k=1}^m \gamma_{j_k} u_{\lambda_{j_k}}\right)}{\left\|K\left(\sum_{k=1}^m \gamma_{j_k} u_{\lambda_{j_k}}\right)\right\|_{\mathcal{H}}^2} K^* K \left(\sum_{k=1}^m \gamma_{j_k} u_{\lambda_{j_k}}\right) \notin \partial J \left(\sum_{k=1}^m \gamma_{j_k} u_{\lambda_{j_k}}\right).$$

Proof. The proof is by contradiction. Hence we assume that $\sum_{k=1}^m \gamma_{j_k} u_{\lambda_{j_k}}$ is a singular vector. Using **(OC)** and Proposition 3.4 we can explicitly compute the singular value:

$$\frac{J\left(\sum_{k=1}^m \gamma_{j_k} u_{\lambda_{j_k}}\right)}{\left\|K\left(\sum_{k=1}^m \gamma_{j_k} u_{\lambda_{j_k}}\right)\right\|_{\mathcal{H}}^2} = \frac{\sum_{k=1}^m \gamma_{j_k} \lambda_{j_k}}{\sum_{k=1}^m \gamma_{j_k}^2}.$$

Using the characterisation (2.6) of the subdifferential we then get

$$\begin{aligned} J(v) &\geq \left\langle \frac{\sum_{k=1}^m \gamma_{j_k} \lambda_{j_k}}{\sum_{k=1}^m \gamma_{j_k}^2} K^* K \left(\sum_{k=1}^m \gamma_{j_k} u_{\lambda_{j_k}}\right), v \right\rangle \\ &= \frac{\sum_{k=1}^m \gamma_{j_k} \lambda_{j_k}}{\sum_{k=1}^m \gamma_{j_k}^2} \sum_{k=1}^m \left\langle \gamma_{j_k} K^* K u_{\lambda_{j_k}}, v \right\rangle. \end{aligned}$$

for all $v \in \text{dom}(J)$. Set $v = u_{\lambda_{j_1}}$. Using $J(u_{\lambda_{j_1}}) = \lambda_{j_1}$ and **(OC)** in the above we obtain

$$\lambda_{j_1} \geq \frac{\sum_{k=1}^m \gamma_{j_k} \lambda_{j_k}}{\sum_{k=1}^m \gamma_{j_k}^2} \gamma_{j_1}.$$

Rearranging and writing out the first term of both sums, the above implies

$$\lambda_{j_1} \gamma_{j_1}^2 + \lambda_{j_1} \sum_{k=2}^m \gamma_{j_k}^2 = \lambda_{j_1} \sum_{k=1}^m \gamma_{j_k}^2 \geq \gamma_{j_1} \sum_{k=1}^m \gamma_{j_k} \lambda_{j_k} = \gamma_{j_1}^2 \lambda_{j_1} + \gamma_{j_1} \sum_{k=2}^m \gamma_{j_k} \lambda_{j_k}.$$

We let the term $\lambda_{j_1} \gamma_{j_1}^2$ on both sides cancel and rearrange again to obtain

$$\frac{\lambda_{j_1}}{\gamma_{j_1}} \geq \frac{\sum_{k=2}^m \gamma_{j_k} \lambda_{j_k}}{\sum_{k=2}^m \gamma_{j_k}^2} = \frac{\gamma_{j_2} \lambda_{j_2} + \sum_{k=3}^m \gamma_{j_k} \lambda_{j_k}}{\sum_{k=2}^m \gamma_{j_k}^2}.$$

Using the ordering $\lambda_{j_k}/\gamma_{j_k} < \lambda_{j_{k+1}}/\gamma_{j_{k+1}}$ then for $k \geq 3$ we get $\lambda_{j_k} > \frac{\lambda_{j_2}}{\gamma_{j_2}} \gamma_{j_k}$. Applying this in the above we reach

$$\frac{\lambda_{j_1}}{\gamma_{j_1}} > \frac{\gamma_{j_2} \lambda_{j_2} + \frac{\lambda_{j_2}}{\gamma_{j_2}} \sum_{k=3}^m \gamma_{j_k}^2}{\sum_{k=2}^m \gamma_{j_k}^2} = \frac{\lambda_{j_2}}{\gamma_{j_2}} \left(\gamma_{j_2}^2 + \sum_{k=3}^m \gamma_{j_k}^2 \right) = \frac{\lambda_{j_2}}{\gamma_{j_2}}.$$

This is a contradiction since by assumption $\lambda_{j_1}/\gamma_{j_1} < \lambda_{j_2}/\gamma_{j_2}$. Hence the assumption that $\sum_{k=1}^m \gamma_{j_k} u_{\lambda_{j_k}}$ is a singular vector is false. \square

4. Decomposition of Arbitrary Input Data

Having addressed the question of when a finite sum of singular vectors can be decomposed via the ISS flow (2.10), we want to draw our attention to a rather converse question: when is the first non-trivial solution of (2.10) a singular vector? We first give a characterisation of the first non-trivial solution of the ISS flow when the data f is given by $f = K\omega$, where ω satisfies the strong source condition (SSC). Then we give a characterisation of the first non-trivial ISS solution for arbitrary data $f \in \mathcal{H}$ and observe that it satisfies the source condition (SC). Finally, we investigate under which conditions this first non-trivial solution is a singular vector. Throughout the section Assumption 2.10 will be used.

4.1. Strong Source Condition

We identify the first step of the ISS flow for given data $f = K\omega$, where $\omega \in \mathcal{U} \setminus \{0\}$ is the source element for the strong source condition (SSC), i.e. $K^*K\omega \in \partial J(u^\dagger)$ for some $u^\dagger \in \text{dom}(J)$. The first non-trivial solution of the ISS flow under this assumption for the data is then characterised via the following lemma.

Lemma 4.1. *Let (SSC) be satisfied, i.e. there exists a source element $\omega \in \mathcal{U} \setminus \{0\}$ such that $K^*K\omega \in \partial J(u^\dagger)$. Let $f = K\omega$ be the given data for the ISS flow (2.10). Then a first non-trivial solution of the flow is given by*

$$u(t_1) = \frac{J(u^\dagger)}{\|Ku^\dagger\|_{\mathcal{H}}^2} u^\dagger, \quad t_1 = 1.$$

Proof. Starting with $u(t) = 0$ for $0 \leq t < t_1$ we need to compute t_1 and $u(t_1)$. By the ISS flow we get

$$\partial_t p(t) = K^*f = K^*K\omega, \quad 0 \leq t < t_1.$$

Using $p(0) = 0$ and integrating $\int_0^t \partial_\tau p(\tau) d\tau$ we get

$$p(t) = tK^*K\omega, \quad 0 \leq t < t_1.$$

We need to show that $p(t) \in \partial J(u(t)) = \partial J(0)$ for $0 \leq t < t_1$. Using the strong source condition and the characterisation (2.6) of the subdifferential, we obtain

$$\langle p(t), v \rangle = t \langle K^*K\omega, v \rangle \leq tJ(v) < t_1 J(v), \quad \forall v \in \mathcal{U}.$$

Hence choosing $t_1 = 1$ we know that $p(t) \in \partial J(0)$ for $0 \leq t < t_1$. At the time t_1 we require $p(t_1) = t_1 K^*K\omega = K^*K\omega \in \partial J(u(t_1))$. Hence choosing $u(t_1) = \gamma u^\dagger$ for $\gamma \geq 0$ the condition is satisfied by the positive scaling invariance (2.7) of the subdifferential. We use the necessary orthogonality condition (2.11) to determine γ :

$$0 = \langle K^*(K\omega - \gamma Ku^\dagger), u^\dagger \rangle = \langle K^*K\omega, u^\dagger \rangle - \gamma \|Ku^\dagger\|_{\mathcal{H}}^2 = J(u^\dagger) - \gamma \|Ku^\dagger\|_{\mathcal{H}}^2,$$

i.e. $\gamma = J(u^\dagger) / \|Ku^\dagger\|_{\mathcal{H}}^2 \geq 0$. Hence a first non-trivial step of the ISS flow is

$$u(t_1) = \frac{J(u^\dagger)}{\|Ku^\dagger\|_{\mathcal{H}}^2} u^\dagger, \quad t_1 = 1.$$

\square

4.2. Source Condition

We emphasise the connection between the first non-trivial solution of the ISS flow (2.10) and the source condition (SC) for arbitrary data f . It turns out that the first solution satisfies the source condition, and under the special kind of source condition (2.20) the first non-trivial solution is also a singular vector.

We show that the first non-trivial solution of the ISS flow satisfies the source condition (SC). Since $u(t) = 0$ for $0 \leq t < t_1$ for some $t_1 > 0$, using the ISS flow we get

$$\partial_t p(t) = K^* f, \quad 0 \leq t < t_1.$$

Using $p(0) = 0$ and integration yields

$$p(t) = tK^* f, \quad 0 \leq t < t_1.$$

At $t = t_1$ we then require

$$p(t_1) = t_1 K^* f \in \partial J(u(t_1))$$

by the definition of the ISS flow. Setting $v = t_1 f \in \mathcal{H}$ we obtain $K^* v \in \partial J(u(t_1))$ which is exactly the source condition. Hence $u(t_1)$ satisfies the source condition.

In Theorem 2.7 it has been shown that for data given as $f = \gamma K u_\lambda + g$, for some $g \in \mathcal{H}$, the first non-trivial solution of the ISS solution at time $t_1 = \lambda\eta/(\lambda + \gamma\eta - \mu) < t_2 = \eta$ is given via $u(t_1) = (\gamma + (\lambda - \mu)/\eta)u_\lambda$ if the condition (2.20) for some $\gamma > \mu/\eta$ is met. We also want to emphasise that (2.20) is in fact nothing but the source condition (SC) for the particular choice $v = \mu K u_\lambda + \eta g$. So we do know from Theorem 2.7 that under this specific kind of source condition the first non-trivial ISS solution is indeed a generalised singular vector. This can be of practical interest if we have arbitrary data f that can be decomposed into a singular vector and an element g such that (2.20) is satisfied. However, one has to be very careful that the conditions for a successful application of Theorem 2.7 are actually met, as the following example shows.

Example 4.2. Let us consider the embedding operator $K = I : \text{BV}([0, 1]) \rightarrow L^2([0, 1])$ and the one-dimensional total variation regularisation

$$J(u) = \text{TV}(u) := \sup_{\substack{\varphi \in C_0^\infty([0, 1]) \\ \|\varphi\|_{L^\infty([0, 1])} \leq 1}} \int_0^1 u(x) \varphi'(x) dx.$$

Given the function $f(x) = -\cos(2\pi x)$ it is quite obvious that we can decompose f into $f = \gamma u_\lambda + g$ for a constant $\gamma > 0$,

$$u_\lambda(x) = \begin{cases} 1, & x \in [\frac{1}{4}, \frac{3}{4}] \\ -1, & \text{else} \end{cases} \quad \text{and} \quad g(x) = \begin{cases} -\gamma - \cos(2\pi x), & x \in [\frac{1}{4}, \frac{3}{4}] \\ \gamma - \cos(2\pi x), & \text{else} \end{cases}.$$

Hence, if we could verify (2.20) for two constants μ and η with $\mu/\eta < \gamma$, we would know from Theorem 2.7 that $(\gamma + (\lambda - \mu)/\eta)u_\lambda$ would be the first non-trivial solution of the ISS flow. In order to do so we have to characterise the subdifferential of TV, which reads as

$$\partial \text{TV}(u) = \{q' \mid \|q\|_{L^\infty([0, 1])} \leq 1, q(0) = q(1) = 0, \text{ and } \langle q', u \rangle = \text{TV}(u)\}.$$

In our case we therefore have to find constants μ and η and a function q with $q' = \mu u_\lambda + \eta g$, $q(0) = q(1) = 0$, $\|q\|_{L^\infty([0,1])} \leq 1$ and $\langle q', u_\lambda \rangle = \text{TV}(u_\lambda)$. We make the assumption $\eta = 2\pi$ and verify that we can find a function q that satisfies all conditions listed above for some μ . We do so by considering $q(x) := -\sin(2\pi x)$. We immediately observe $q(0) = q(1) = 0$ and $\|q\|_{L^\infty([0,1])} = 1$. For the choice $\eta = 2\pi$ we further compute

$$q'(x) = -2\pi \cos(2\pi x) = \mu u_\lambda(x) + \eta \begin{cases} -\frac{\mu}{\eta} - \cos(2\pi x), & x \in [\frac{1}{4}, \frac{3}{4}] \\ \frac{\mu}{\eta} - \cos(2\pi x), & \text{else} \end{cases}.$$

For the choice $\mu = \gamma\eta$ we then obtain $-q' = \mu u_\lambda + \eta w$, and further observe $\langle q', u_\lambda \rangle = 4 = \text{TV}(u_\lambda)$. Hence, we have verified $\mu u_\lambda + \eta w \in \partial \text{TV}(u_\lambda)$ for $\eta = 2\pi$ and $\mu = \gamma\eta$. Unfortunately, we do not obtain $\mu/\eta < \gamma$, which means that we cannot apply Theorem 2.7 to automatically conclude that the first non-trivial solution of the ISS flow is given via a singular vector, though it seems tempting at first glance.

4.3. Arbitrary input data

In this section we want to characterise the first non-trivial solution of the ISS flow for arbitrary given data. We investigate under which conditions the first solution is guaranteed to be a singular vector. This investigation leads to the definition of dual singular vectors in the data space.

Under the norm-inequality assumption (3.2) for J we can characterise the first non-trivial solution of the ISS flow (2.10) via the following lemma.

Lemma 4.3 (First Non-trivial Solution of ISS Flow). *Assume that J satisfies the norm-inequality assumption (3.2). Let $f \in \mathcal{H} \setminus \ker(K^*)$ be arbitrary given data for the ISS flow (2.10) satisfying $\langle K^*f, v \rangle = 0$ for all $v \in \ker(J)$. Then the first non-trivial solution of the flow is characterised by*

$$u(t_1) = c_1 u_1$$

where $t_1 = c_0 / \|K^*f\|_{\mathcal{U}^*}$ (c_0 being the norm-inequality constant), $c_1 = J(u_1)/t_1$ and $u_1 = v_1 / \|Kv_1\|_{\mathcal{H}}$ with

$$v_1 \in \arg \min_{v \in \ker(J)^\perp} \{J(v) - t_1 \langle K^*f, v \rangle\}. \quad (4.1)$$

Proof. For $0 \leq t < t_1$ we have $u(t) = 0$. The ISS flow then gives

$$\partial_t p(t) = K^*f, \quad 0 \leq t < t_1.$$

Using $p(0) = 0$ and integration we obtain $p(t) = tK^*f$ for $0 \leq t < t_1$. This implies

$$\|p(t)\|_{\mathcal{U}^*} = t \|K^*f\|_{\mathcal{U}^*} < t_1 \|K^*f\|_{\mathcal{U}^*} = c_0.$$

Together with the assumption that $\langle K^*f, v \rangle = 0$ for all $v \in \ker(J)$, Proposition 3.11 then implies $p(t) \in \partial J(0)$. Hence the ISS flow is satisfied for $0 \leq t < t_1$. At $t = t_1$ we need $p(t_1) = t_1 K^*f \in \partial J(c_1 u_1) = \partial J(u_1) = \partial J(v_1)$ for $c_1 u_1$ to be the solution at time t_1 . This is exactly the optimality condition of

$$v_1 \in \arg \min_{v \in \ker(J)^\perp} \{J(v) - t_1 \langle K^*f, v \rangle\},$$

since the Lagrange multiplier for the constraint $v \in \ker(J)^\perp$ is zero following the same procedure as of that for the generalised singular vectors in Section 2. Finally, we require $u(t_1)$ to satisfy the necessary orthogonality condition (2.11):

$$0 = \langle K^*(f - Ku(t_1)), u(t_1) \rangle = c_1 \langle K^*f, u_1 \rangle - c_1^2 = c_1 \left(\frac{J(u_1)}{t_1} - c_1 \right).$$

We have used that $t_1 K^*f \in \partial J(u_1)$ for the last equality. From the above we obtain

$$c_1 = \frac{J(u_1)}{t_1}.$$

□

The interesting question now is when the first ISS solution u_1 characterised via (4.1) is a singular vector. We need to investigate if $J(u_1)K^*Ku_1 \in \partial J(u_1)$. Using the characterisation of the subdifferential (2.6) we need to show

$$\langle J(u_1)K^*Ku_1, u_1 \rangle = J(u_1) \quad \text{and} \quad J(u_1)K^*Ku_1 \in \partial J(0).$$

The first condition is easy to verify since $\|Ku_1\|_{\mathcal{H}} = 1$. For the second condition we use Proposition 3.11, thus we need $\|J(u_1)K^*Ku_1\|_{\mathcal{U}^*} \leq c_0$. For $J(u_1)$ we use the estimate

$$J(u_1) = t_1 \langle K^*f, u_1 \rangle = \frac{c_0}{\|K^*f\|_{\mathcal{U}^*}} \langle f, Ku_1 \rangle \leq \frac{c_0}{\|K^*f\|_{\mathcal{U}^*}} \|f\|_{\mathcal{H}} \|Ku_1\|_{\mathcal{H}} = c_0 \frac{\|f\|_{\mathcal{H}}}{\|K^*f\|_{\mathcal{U}^*}}.$$

Then we obtain

$$\|J(u_1)K^*Ku_1\|_{\mathcal{U}^*} \leq c_0 \frac{\|f\|_{\mathcal{H}}}{\|K^*f\|_{\mathcal{U}^*}} \|K^*\| \|Ku_1\|_{\mathcal{H}} = c_0 \frac{\|f\|_{\mathcal{H}} \|K^*\|}{\|K^*f\|_{\mathcal{U}^*}},$$

where $\|K^*\|$ is the standard operator norm

$$\|K^*\| = \sup_{\omega \in \mathcal{H} \setminus \{0\}} \frac{\|K^*\omega\|_{\mathcal{U}^*}}{\|\omega\|_{\mathcal{H}}}.$$

Hence we see that if $\|K^*f\|_{\mathcal{U}^*} = \|K^*\| \|f\|_{\mathcal{H}}$ then the first ISS solution u_1 is actually a singular vector of J . We note here that one should be very careful using the correct norms. To emphasise this we consider the following example:

Example 4.4. Let $K = I : H_0^1(\Omega) \rightarrow L_0^2(\Omega)$ be an imbedding operator and Ω a bounded domain in \mathbb{R}^2 . Here subscript zero means that the functions are zero on the boundary of Ω . Let $J(u) = \|\nabla u\|_{L^2(\Omega)}$. Using Poincaré's inequality $\|u\|_{L^2(\Omega)} \leq C \|\nabla u\|_{L^2(\Omega)}$ for some $C > 0$ and $u \in \ker(J)^\perp$, we can verify that J satisfies the norm-inequality assumption (3.2) for $c_0 = 1/(1 + C^2)^{1/2}$. Let $f \in H^1(\Omega)$ be the data for the ISS flow. If f satisfies

$$\|I^*f\|_{H^1(\Omega)} = \|I^*\| \|f\|_{L^2(\Omega)} = \|f\|_{L^2(\Omega)},$$

the first non-trivial solution of the ISS flow is a singular vector. The equality is not always satisfied even though K is an embedding operator. We see that we need

$$\left(\|f\|_{L^2(\Omega)}^2 + \|\nabla f\|_{L^2(\Omega)}^2 \right)^{1/2} = \|f\|_{L^2(\Omega)} \quad \Leftrightarrow \quad \|\nabla f\|_{L^2(\Omega)} = 0.$$

The condition $\|K^*f\|_{\mathcal{U}^*} = \|K^*\| \|f\|_{\mathcal{H}}$ gives rise to the following theorem:

Theorem 4.5. *Assume that J satisfies the norm-inequality assumption (3.2). If $f \in \mathcal{H} \setminus \ker(K^*)$ is a dual singular vector, i.e.*

$$\frac{\|K^*f\|_{\mathcal{U}^*}}{\|f\|_{\mathcal{H}}^2} f \in \partial \|K^*f\|_{\mathcal{U}^*}, \quad (4.2)$$

*and satisfies $\langle K^*f, v \rangle = 0$ for all $v \in \ker(J)$, then the first non-trivial solution u_1 of the ISS flow (2.10) characterised via (4.1) is a (primal) singular vector of J .*

Proof. We have shown above that if f satisfies $\|K^*f\|_{\mathcal{U}^*} = \|K^*\| \|f\|_{\mathcal{H}}$, then u_1 is a singular vector. Since

$$\|K^*\| = \sup_{\omega \in \mathcal{H}} \frac{\|K^*\omega\|_{\mathcal{U}^*}}{\|\omega\|_{\mathcal{H}}}$$

we see that in order for $\|K^*f\|_{\mathcal{U}^*} = \|K^*\| \|f\|_{\mathcal{H}}$ to hold true, f has to satisfy

$$f \in \arg \max_{\omega \in \mathcal{H}} \frac{\|K^*\omega\|_{\mathcal{U}^*}}{\|\omega\|_{\mathcal{H}}}.$$

The optimality condition for this is

$$0 \in \frac{\partial \|K^*f\|_{\mathcal{U}^*} \|f\|_{\mathcal{H}} - \|K^*f\|_{\mathcal{U}^*} \frac{f}{\|f\|_{\mathcal{H}}}}{\|f\|_{\mathcal{H}}^2}$$

which implies

$$\frac{\|K^*f\|_{\mathcal{U}^*}}{\|f\|_{\mathcal{H}}^2} f \in \partial \|K^*f\|_{\mathcal{U}^*}.$$

□

Remark 4.6. Setting $J_2(\omega) = \|K^*\omega\|_{\mathcal{U}^*}$, $\omega \in \mathcal{H}$, we see from the above that f has to be a singular vector of J_2 (and $I : \mathcal{H} \rightarrow \mathcal{H}$). J_2 is absolutely one-homogeneous and hence we can reuse all the results for absolutely one-homogeneous functionals.

5. Numerical Results

We conclude the paper with numerical results that demonstrate the sufficiency of the **(SUB0)** condition when we wish to decompose a linear combination of generalised singular vectors. In order to verify the theoretical results provided in Section 3, we need to be able to solve the ISS flow (2.10) numerically. In [69] it has been shown that solutions of (2.10) can be computed directly in case of $J(u) = \|u\|_{\ell^1(\mathbb{R}^m)}$ via what has been termed the adaptive inverse scale space method (aISS). In [70] the aISS has been extended to general polyhedral functionals J and could therefore also be used to numerically verify the behaviour of functionals like the anisotropic total variation, for example. For arbitrary convex, absolutely one-homogeneous functionals J , however, there does not exist a way to efficiently compute the exact ISS solution at each time step numerically to the best of our knowledge. Therefore, for simplicity, we are going to restrict ourselves to the case of $J(u) = \|u\|_{\ell^1(\mathbb{R}^n)}$ for the numerical results. We compute the inverse scale space solution for a composition satisfying the **(SUB0)** condition. We also demonstrate the need for the **(SUB0)** condition computing ISS solutions of compositions following Examples 3.8 and 3.18.

5.1. (SUB0) Observance

With our first example we simply verify that the result of Theorem 3.14 is observed in practice. In order to do so, we consider the operator $K : \ell^1(\mathbb{R}^8) \supseteq \ell^2(\mathbb{R}^8) \rightarrow \ell^2(\mathbb{R}^8)$ defined by $(Ku)_k = \sum_{j=-1}^1 u_{k-j}g_j$ for $g = (1, 1, 0)^T/\sqrt{2}$ from Example 3.8, respectively Example 3.18, the absolutely one-homogeneous functional $J(u) = \|u\|_{\ell^1(\mathbb{R}^8)}$, and the following three K -normalised and K -orthogonal singular vectors

$$\begin{aligned} u_{\lambda_1} &= (0, 0, 0, 1, -1, 0, 0, 0)^T, \\ u_{\lambda_2} &= (0, -1, 0, 0, 0, 0, 0, 0)^T, \\ u_{\lambda_3} &= (0, 0, 0, 0, 0, 0, 1, 0)^T, \end{aligned}$$

with $\lambda_1 = 2$ and $\lambda_2 = \lambda_3 = 1$. We now verify that these singular vectors satisfy the (SUB0) condition in the order they are defined, i.e. we have to verify $\lambda_1 K^* K u_{\lambda_1} \in \partial J(0)$, $\lambda_1 K^* K u_{\lambda_1} + \lambda_2 K^* K u_{\lambda_2} \in \partial J(0)$ and $\lambda_1 K^* K u_{\lambda_1} + \lambda_2 K^* K u_{\lambda_2} + \lambda_3 K^* K u_{\lambda_3} \in \partial J(0)$. The first condition trivially follows from u_{λ_1} being a generalised singular vector, the second and third follow from

$$\begin{aligned} \sum_{j=1}^2 \lambda_j K^* K u_{\lambda_j} &= (-1/2, -1, 1/2, 1, -1, -1, 0, 0)^T \in \partial J(0), \\ \sum_{j=1}^3 \lambda_j K^* K u_{\lambda_j} &= (-1/2, -1, 1/2, 1, -1, -1/2, 1, 1/2)^T \in \partial J(0). \end{aligned}$$

Hence, (SUB0) is satisfied and Theorem 3.14 guarantees that (2.10) and therefore the aISS algorithm will decompose any composition $f = \sum_{j=1}^3 \gamma_j K u_{\lambda_j}$ with $\lambda_1/\gamma_1 < \lambda_2/\gamma_2 < \lambda_3/\gamma_3$ into $\gamma_1 u_{\lambda_1}$, $\gamma_1 u_{\lambda_1} + \gamma_2 u_{\lambda_2}$ and $\gamma_1 u_{\lambda_1} + \gamma_2 u_{\lambda_2} + \gamma_3 u_{\lambda_3}$. Figure 5 shows the three iterations of the aISS method until convergence, for $f = \sum_{j=1}^3 \gamma_j K u_{\lambda_j}$ with $\gamma_1 = 5$, $\gamma_2 = 2$ and $\gamma_3 = 1$. It can be seen that the aISS algorithm and therefore (2.10) indeed decomposes the signal f as stated by Theorem 3.14.

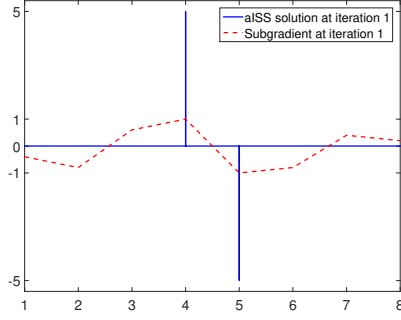
5.2. (SUB0) Violations

Following Example 3.8 we also want to verify numerically that the absence of the (SUB0) condition in fact can lead to results different from the decomposition of the data into the original singular vectors. We use the operator K , the functional J , and the two singular vectors

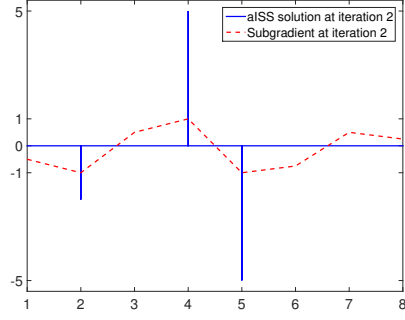
$$u_{\lambda_1} = (0, 0, 1, 0, 0)^T \quad \text{and} \quad u_{\lambda_2} = \frac{1}{\sqrt{2}}(0, 1, 0, -1, 0)^T,$$

from example 3.8. We know that the two singular vectors are K -orthogonal but violate the (SUB0) condition. We can then for instance define $f = K u_{\lambda_1} + \frac{3}{2\sqrt{2}} K u_{\lambda_2}$ as potential input data that we want to decompose. Note that we have ensured $\lambda_1/\gamma_1 = 1 < 4/3 = \lambda_2/\gamma_2$. In order to compute a solution of the ISS flow (2.10) we again use the aISS method. The first non-trivial solution $u(t_1)$ of the aISS method reads as

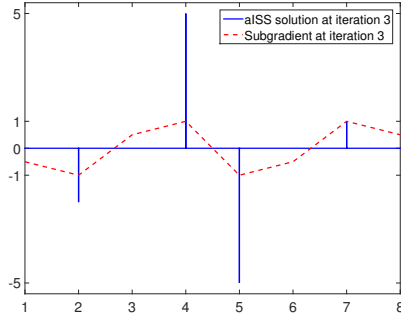
$$u(t_1) = (0, 5/4, 0, 0, 0)^T.$$



(a) First aISS solution and subgradient.



(b) Second aISS solution and subgradient.



(c) Third and final aISS solution and subgradient.

Figure 5: The three iterations of the aISS algorithm until convergence for the example of Section 5.1. Figure 5a shows both the primal and the dual variable, i.e. $u(t_1)$, $p(t_1)$, at $t = t_1$, whereas Figure 5b and 5c show the same for t_2 and t_3 , respectively. It is easily observed that the aISS algorithm has decomposed the signal f into the original singular vectors with no loss of contrast, as stated by Theorem 3.14.

This solution is obviously as sparse as u_{λ_1} and also a singular vector (with the same singular value after normalisation) but corresponds neither to u_{λ_1} nor to u_{λ_2} . More importantly, it is a better fit to the data f with respect to the ℓ^2 -norm, as we have $\|Ku(t_1) - f\|_{\ell^2(\mathbb{R}^m)}^2 = 9/16$ as opposed to $\|Ku_{\lambda_1} - f\|_{\ell^2(\mathbb{R}^m)}^2 = 9/8$. Hence, it is logical that the first ISS solution is not u_{λ_1} . The remaining ISS solutions are

$$u(t_2) = (0, 1, 1/2, 0, 0)^T \quad \text{and} \quad u(t_3) = (0, 3/4, 1, -3/4, 0)^T,$$

where $u(t_3)$ satisfies $u(t_3) = u_{\lambda_1} + \frac{3}{2\sqrt{2}}u_{\lambda_2}$. What is interesting about this example is that the ISS flow (2.10), whilst converging to a J -minimal solution of $K^*Ku = K^*f$, most importantly has to minimise the residual in each step. This dominates the behaviour of keeping J minimal at the same time, as we observe $J(u(t_1)) > J(u_{\lambda_1})$ but $\|Ku(t_1) - f\|_{\ell^2(\mathbb{R}^m)} < \|Ku_{\lambda_1} - f\|_{\ell^2(\mathbb{R}^m)}$.

We can obviously apply aISS method to compute solutions for compositions made from the singular vectors described in Example 3.18 as well. We pick $\gamma_1 = 9$, $\gamma_2 = 8$,

$\gamma_3 = 3\sqrt{2}$, $\gamma_4 = 2$, and $\gamma_5 = 1$ in order to have $\lambda_k/\gamma_k < \lambda_{k+1}/\gamma_{k+1}$ for $k = 1, \dots, 4$. In this case the first non-trivial solution $u(t_1)$ of the aISS method is given by

$$u(t_1) = (0, 0, 0, 0, -11/2, 0, 0, 0, 0)^T.$$

The recovered peak does not match any of the singular vectors u_{λ_j} , $j = 1, \dots, 5$. This shows that in order for the ISS flow to give a decomposition into the singular vectors of f it is not enough that the full sum of the subgradients is in the subdifferential of J at zero. We need the partial sums as well. The remaining solutions are

$$\begin{aligned} u(t_2) &= (0, 0, 5, 0, -11/2, 0, 5, 0, 0)^T, \\ u(t_3) &= (0, 0, 5/4, 15/2, -37/4, 0, 5, 0, 0)^T, \\ u(t_4) &= (0, 0, 0, 12, -17, 11, 0, 0, 0)^T = \gamma_1 u_{\lambda_1} + \gamma_2 u_{\lambda_2} + \gamma_3 u_{\lambda_3}, \\ u(t_5) &= (0, -2, 0, 12, -17, 11, 0, 0, 0)^T = \gamma_1 u_{\lambda_1} + \gamma_2 u_{\lambda_2} + \gamma_3 u_{\lambda_3} + \gamma_4 u_{\lambda_4}, \\ u(t_6) &= (0, -2, 0, 12, -17, 11, 0, -1, 0)^T = \gamma_1 u_{\lambda_1} + \gamma_2 u_{\lambda_2} + \gamma_3 u_{\lambda_3} + \gamma_4 u_{\lambda_4} + \gamma_5 u_{\lambda_5}. \end{aligned}$$

The solutions at steps two and three cannot be written as any linear combination of the five singular vectors u_{λ_j} , $j = 1, \dots, 5$. From step four the solutions can be written as linear combinations of the singular vectors with no loss of contrast. However, we are not able to separate u_{λ_1} , u_{λ_2} , and u_{λ_3} .

6. Conclusions & Outlook

In this paper we have investigated the possibility of using the inverse scale space flow as a decomposition method for absolutely one-homogeneous regularisation functionals. We have formulated conditions under which the inverse scale space flow will give a decomposition of a linear combination of generalised singular vectors. Furthermore, we have investigated the behaviour of the flow for arbitrary data. We have given a characterisation of the first non-trivial solution and shown that this solution is a primal singular vector when the input data is what we have named a dual singular vector.

For the inverse scale space decomposition we require the generalised singular vectors representing the data to be orthogonal when mapped to the data space by the forward operator. This condition, however, is quite restrictive and in future research we aim to relax the condition. The idea is that the inverse scale space flow will then still give a decomposition into the singular vectors but with time-varying coefficients. Another subject of interest is to investigate whether the data can always be decomposed into a linear combination of singular vectors and a remainder that can be seen as noise satisfying special conditions such as (2.20). Extending the decomposition results to data containing noise in the same manner as in [3, Section 7.2] we then know that the inverse scale space flow will always give a decomposition into singular vectors.

Another direction for future research is the numerical computation of inverse scale space solutions as well as the numerical computation of generalised singular vectors. As mentioned earlier, the former is currently restricted to polyhedral regularisation functionals (cf. [69, 70]). For the latter there has been made substantial progress in [71, 72, 73, 74], but many open questions need yet to be addressed. How can we compute ground states as well as non-trivial singular vectors with larger singular

values? How can we incorporate (linear) forward operators? How can we compute dual singular vectors in a robust fashion? These are only few of many open questions that need to be addressed in future research.

Acknowledgments

MFS acknowledges support from the Advanced Grant No. 291405 from the European Research Council as part of the project HD-Tomo. MB and CBS acknowledge support from the EPSRC Grant EP/M00483X/1, the Leverhulme Trust Grant 'Breaking the non-convexity barrier'. MB further acknowledges support from the Isaac Newton Trust and the Leverhulme Trust Early Career Fellowship Grant 'Learning from mistakes: a supervised feedback-loop for imaging applications'. CBS acknowledges support from the EPSRC Centre 'EP/N014588/1' and the Cantab Capital Institute for the Mathematics of Information. The authors thank Martin Burger, Antonin Chambolle, Guy Gilboa and Michael Möller for fruitful discussions.

Data Statement

The corresponding ©MATLAB-code will be put into a repository and made publicly available as the final version is submitted.

References

- [1] Stewart G W 1993 *SIAM review* **35** 551–566
- [2] Engl H W, Hanke M and Neubauer A 1996 *Regularization of inverse problems* vol 375 (Springer Science & Business Media)
- [3] Benning M and Burger M 2013 *Methods and Applications of Analysis* **20** 295–334
- [4] Agueh M 2010 *Advances in Calculus of Variations* **3** 233–261
- [5] Berestycki H and Lions P 1980 Existence of stationary states in nonlinear scalar field equations *Bifurcation phenomena in mathematical physics and related topics* (Springer) pp 269–292
- [6] Browder F E *et al.* 1965 *Bull. Amer. Math. Soc* **71** 176–183
- [7] Dacorogna B, Gangbo W and Subia N 1992 Sur une généralisation de l'inégalité de wirtinger *Annales de l'IHP Analyse non linéaire* vol 9 pp 29–50
- [8] Dolbeault J, Esteban M J and Loss M 2012 *Journal of Mathematical Physics* **53** 095204
- [9] Franzina G and Lamberti P D 2010 *Electron. J. Differential Equations* **26**
- [10] Henrot A 2006 *Extremum problems for eigenvalues of elliptic operators* (Springer Science & Business Media)
- [11] Kawohl B 2000 *Discrete and continuous dynamical systems* **6** 683–690
- [12] Kawohl B and Fridman V 2003 *Comment. Math. Univ. Carolin* **44** 659–667
- [13] Kawohl B, Lucia M, Prashanth S *et al.* 2007 *Advances in Differential Equations* **12** 407–434
- [14] Weinstein M I 1983 *Communications in Mathematical Physics* **87** 567–576
- [15] Fujimoto K 2004 What are singular values of nonlinear operators? *Decision and Control, 2004. CDC. 43rd IEEE Conference on* vol 2 (IEEE) pp 1623–1628
- [16] Meyer Y 2001 *Oscillating patterns in image processing and nonlinear evolution equations: the fifteenth Dean Jacqueline B. Lewis memorial lectures* vol 22 (American Mathematical Soc.)
- [17] Alter F, Caselles V and Chambolle A 2005 *Mathematische Annalen* **332** 329–366
- [18] Alter F, Caselles V and Chambolle A 2005 *Interfaces and Free Boundaries* **7** 29–53
- [19] Caselles V, Chambolle A, Moll S and Novaga M 2008 A characterization of convex calibrable sets in with respect to anisotropic norms *Annales de l'IHP Analyse non linéaire* vol 25 pp 803–832
- [20] Edelman A, Arias T A and Smith S T 1998 *SIAM journal on Matrix Analysis and Applications* **20** 303–353
- [21] Chambolle A, Duval V, Peyré G and Poon C 2016 *arXiv preprint arXiv:1602.00087*
- [22] Chambolle A, Duval V, Peyré G and Poon C 2016 Total variation denoising and support localization of the gradient *Journal of Physics: Conference Series* vol 756 (IOP Publishing) p 012007

- [23] Bühler T and Hein M 2009 Spectral clustering based on the graph p-laplacian *Proceedings of the 26th Annual International Conference on Machine Learning (ACM)* pp 81–88
- [24] Bresson X and Szlam A D 2010 Total variation, cheeger cuts *Proceedings of the 27th International Conference on Machine Learning (ICML-10)* pp 1039–1046
- [25] Rudin L I, Osher S and Fatemi E 1992 *Physica D: Nonlinear Phenomena* **60** 259–268
- [26] Candes E J and Tao T 2005 *IEEE transactions on information theory* **51** 4203–4215
- [27] Candes E J and Tao T 2006 *IEEE transactions on information theory* **52** 5406–5425
- [28] Donoho D L 2006 *IEEE Transactions on information theory* **52** 1289–1306
- [29] Donoho D L and Elad M 2003 *Proceedings of the National Academy of Sciences* **100** 2197–2202
- [30] Bredies K, Kunisch K and Pock T 2010 *SIAM Journal on Imaging Sciences* **3** 492–526
- [31] Setzer S and Steidl G 2008 *Approximation* **12** 360–386
- [32] Setzer S, Steidl G and Teuber T 2011 *Comm. Math. Sci* **9** 797–872
- [33] Bredies K and Valkonen T 2011 Inverse problems with second-order total generalized variation constraints *Proceedings of the 9th International Conference on Sampling Theory and Applications (SampTA) 2011, Singapore* vol 2011
- [34] Benning M, Brune C, Burger M and Müller J 2013 *Journal of Scientific Computing* **54** 269–310
- [35] Kutyniok G *et al.* 2012 *Shearlets: Multiscale analysis for multivariate data* (Springer Science & Business Media)
- [36] Goldluecke B, Strelakovsky E and Cremers D 2012 *SIAM Journal on Imaging Sciences* **5** 537–563
- [37] Duran J, Moeller M, Sbert C and Cremers D 2015 A novel framework for nonlocal vectorial total variation based on $\ell^{p,q,r}$ -norms *International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition* (Springer) pp 141–154
- [38] Duran J, Moeller M, Sbert C and Cremers D 2016 *SIAM Journal on Imaging Sciences* **9** 116–151
- [39] Brox T, Weickert J, Burgeth B and Mrázek P 2006 *Image and Vision Computing* **24** 41–55
- [40] Grasmair M and Lenzen F 2010 *Applied Mathematics & Optimization* **62** 323–339
- [41] Lefkimiatis S, Roussos A, Unser M and Maragos P 2013 Convex generalizations of total variation based on the structure tensor with applications to inverse problems *International Conference on Scale Space and Variational Methods in Computer Vision* (Springer) pp 48–60
- [42] Lefkimiatis S, Roussos A, Maragos P and Unser M 2015 *SIAM Journal on Imaging Sciences* **8** 1090–1122
- [43] Estellers V, Soatto S and Bresson X 2015 *IEEE Transactions on Image Processing* **24** 1777–1790
- [44] Brinkmann E M, Burger M and Grah J 2015 Regularization with sparse vector fields: From image compression to tv-type reconstruction *International Conference on Scale Space and Variational Methods in Computer Vision* (Springer) pp 191–202
- [45] Brinkmann E M, Burger M, Rasch J and Sutour C 2016 *arXiv preprint arXiv:1606.05113*
- [46] Cai J F, Chan R H and Shen Z 2008 *Applied and Computational Harmonic Analysis* **24** 131–149
- [47] Cai J F, Osher S and Shen Z 2009 *Multiscale modeling & simulation* **8** 337–369
- [48] Cai J F, Dong B, Osher S and Shen Z 2012 *Journal of the American Mathematical Society* **25** 1033–1089
- [49] Burger M, Gilboa G, Osher S, Xu J *et al.* 2006 *Communications in Mathematical Sciences* **4** 179–212
- [50] Osher S, Burger M, Goldfarb D, Xu J and Yin W 2005 *Multiscale Modeling & Simulation* **4** 460–489
- [51] Zhang X, Burger M and Osher S 2011 *Journal of Scientific Computing* **46** 20–46
- [52] Yin W, Osher S, Goldfarb D and Darbon J 2008 *SIAM Journal on Imaging sciences* **1** 143–168
- [53] Cai J F, Osher S and Shen Z 2009 *Mathematics of Computation* **78** 1515–1536
- [54] Cai J F, Osher S and Shen Z 2009 *Mathematics of Computation* **78** 2127–2136
- [55] Yin W 2010 *SIAM Journal on Imaging Sciences* **3** 856–877
- [56] Moeller M, Brinkmann E M, Burger M and Seybold T 2014 *SIAM Journal on Imaging Sciences* **7** 2771–2806
- [57] Zeune L, van Dalum G, Terstappen L W, van Gils S and Brune C 2016 *arXiv preprint arXiv:1604.06665*
- [58] Gilboa G 2013 *SSVM* **7893** 36–47
- [59] Gilboa G 2014 *SIAM Journal on Imaging Sciences* **7** 1937–1961
- [60] Burger M, Eckardt L, Gilboa G and Moeller M 2015 Spectral representations of one-homogeneous functionals *Scale Space and Variational Methods in Computer Vision* (Springer) pp 16–27
- [61] Horesh D and Gilboa G 2015 Multiscale texture orientation analysis using spectral total-variation decomposition *Scale Space and Variational Methods in Computer Vision* (Springer) pp 486–497
- [62] Gilboa G 2015 *arXiv preprint arXiv:1511.04685*

- [63] Burger M, Gilboa G, Moeller M, Eckardt L and Cremers D 2016 *arXiv preprint arXiv:1601.02912*
- [64] Burger M and Osher S 2013 A guide to the tv zoo *Level Set and PDE Based Reconstruction Methods in Imaging* (Springer) pp 1–70
- [65] Bregman L M 1967 *USSR computational mathematics and mathematical physics* **7** 200–217
- [66] Burger M 2016 Bregman distances in inverse problems and partial differential equations *Advances in Mathematical Modeling, Optimization and Optimal Control* (Springer) pp 3–33
- [67] Showalter D 1967 *Proceedings of the American Mathematical Society* **18** 584–586
- [68] Resmerita E 2005 *Inverse Problems* **21** 1303
- [69] Burger M, Möller M, Benning M and Osher S 2013 *Mathematics of Computation* **82** 269–299
- [70] Moeller M and Burger M 2013 *SIAM Journal on Optimization* **23** 1424–1456
- [71] Hein M and Bühler T 2010 An inverse power method for nonlinear eigenproblems with applications in 1-spectral clustering and sparse pca *Advances in Neural Information Processing Systems* pp 847–855
- [72] Bresson X, Laurent T, Uminsky D and Brecht J V 2012 Convergence and energy landscape for cheeger cut clustering *Advances in Neural Information Processing Systems* pp 1385–1393
- [73] Benning M, Gilboa G and Schönlieb C B 2016 *PAMM* **16** 933–936
- [74] Nosseck R Z and Gilboa G 2016 *arXiv preprint arXiv:1609.08438*